

Tue Jun 22 11:04:04 2004

us-10-018-672-2.rapb

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

CM protein - protein search, using sw model

Run on: June 16, 2004, 11:08:12 ; Search time 49 Seconds
(without alignments)
1586.865 Million cell updates/sec

Title: US-10-018-672-2

Sequence: 1 MAFKINGICALAGIALAG.....TDEVEAKKQFDGVIKW 276

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1158786 seqs, 281726120 residues

Total number of hits satisfying chosen parameters: 1158786

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: Published Applications RA: *
2: /cgn2_6/ptodata/2/pubppa/PCR_NEM_PUB_PEP: *
3: /cgn2_6/ptodata/2/pubppa/US06_NEM_PUB_PEP: *
4: /cgn2_6/ptodata/2/pubppa/US06_PUBCOMB_PEP: *
5: /cgn2_6/ptodata/2/pubppa/US07_NEM_PUB_PEP: *
6: /cgn2_6/ptodata/2/pubppa/US07_PUBCOMB_PEP: *
7: /cgn2_6/ptodata/2/pubppa/US08_NEM_PUB_PEP: *
8: /cgn2_6/ptodata/2/pubppa/US08_PUBCOMB_PEP: *
9: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMB_PEP: *
10: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMB_PEP: *
11: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMB_PEP: *
12: /cgn2_6/ptodata/2/pubppa/US09_PUB_PEP: *
13: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMB_PEP: *
14: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMB_PEP: *
15: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMB_PEP: *
16: /cgn2_6/ptodata/2/pubppa/US10_PUB_PEP: *
17: /cgn2_6/ptodata/2/pubppa/US60_NEM_PUB_PEP: *
18: /cgn2_6/ptodata/2/pubppa/US60_PUBCOMB_PEP: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1368	98.9	276	US-10-282-122A-63386	Sequence 63386, A
2	1007.5	72.8	273	US-10-380-817-2	Sequence 2, Appl1
3	1007.5	72.8	273	US-10-380-817-4	Sequence 8, Appl1
4	1007.5	72.8	273	US-10-380-817-8	Sequence 4, Appl1
5	1007.5	72.8	273	US-10-380-817-10	Sequence 10, Appl1
6	1006.5	72.8	273	US-09-815-242-11067	Sequence 11067, A
7	1006.5	72.8	273	US-10-282-122A-58240	Sequence 58240, A
8	1006.5	72.8	273	US-10-380-817-6	Sequence 6, Appl1
9	888	64.2	276	US-10-282-122A-67387	Sequence 67387, A
10	793.5	57.4	271	US-10-282-122A-76195	Sequence 76195, A
11	782.5	56.6	271	US-09-815-242-10050	Sequence 10050, A
12	782.5	56.6	271	US-10-282-122A-43291	Sequence 43291, A
13	780.5	55.4	271	US-10-282-122A-55826	Sequence 55826, A
14	775.5	54.1	271	US-10-282-122A-59600	Sequence 59600, A
15	757.5	54.8	271	US-10-282-122A-68391	Sequence 68391, A

16	757.5	54.8	271	US-10-282-122A-78174	Sequence 78174, A
17	751.5	54.3	275	US-10-282-122A-77174	Sequence 77174, A
18	743.5	53.8	240	US-09-815-242-11656	Sequence 11656, A
19	562	40.6	295	US-10-282-122A-47704	Sequence 47704, A
20	551	39.8	262	US-10-282-122A-51371	Sequence 51371, A
21	549.5	39.7	272	US-10-282-122A-50375	Sequence 50375, A
22	545	39.4	256	US-10-282-122A-61127	Sequence 61127, A
23	540	39.0	268	US-10-282-122A-49256	Sequence 49256, A
24	527.5	38.1	270	US-10-282-122A-49339	Sequence 49339, A
25	526.5	38.1	265	US-10-282-122A-51201	Sequence 51201, A
26	506.5	36.6	261	US-10-282-122A-67692	Sequence 67692, A
27	502.5	36.3	260	US-09-815-242-12110	Sequence 12110, A
28	502.5	36.3	260	US-10-282-122A-66844	Sequence 66844, A
29	497.5	36.0	259	US-09-815-242-5146	Sequence 5146, Ap
30	497.5	36.0	259	US-10-282-122A-43469	Sequence 43469, A
31	494	35.7	263	US-10-282-122A-52976	Sequence 52976, A
32	488	35.3	256	US-10-282-122A-67627	Sequence 67627, A
33	487	35.2	257	US-10-282-122A-69359	Sequence 69359, A
34	481	34.8	277	US-10-282-122A-57209	Sequence 57209, A
35	481	34.8	277	US-10-282-122A-57630	Sequence 57630, A
36	477.5	34.5	257	US-10-282-122A-54418	Sequence 54418, A
37	472.5	34.2	270	US-10-282-122A-54496	Sequence 54496, A
38	472	34.1	270	US-10-282-122A-52838	Sequence 52838, A
39	468	33.8	256	US-10-282-122A-54417	Sequence 54417, A
40	467.5	33.8	271	US-09-815-242-11460	Sequence 11460, A
41	467.5	33.8	271	US-10-282-122A-59034	Sequence 59034, A
42	462.5	33.4	271	US-09-815-242-11624	Sequence 11624, A
43	462.5	33.4	271	US-10-335-977-4882	Sequence 4882, Ap
44	462.5	33.4	271	US-10-335-977-4883	Sequence 4883, Ap
45	462.5	33.4	273	US-10-282-122A-60558	Sequence 60558, A

RESULT 1
US-10-282-122A-63386
Sequence 63386, A
Publication No. US2004002912391

GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haeselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: EPIRA.034A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: US/10/282,122A
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636

check pricing for New 102 (e) Jul 01/05 101

; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 63386
 ; LENGTH: 276
 ; TYPE: PRF
 ; ORGANISM: Moraxella catarrhalis
 US-10-282-122A-63386

Query Match 98.9%; Score 1368; DB 12; Length 276;
 Best Local Similarity 99.3%; Pred. No. 1,8e-116;
 Matches 274; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MNFGKINGICALASGIALAGCSNOSNEPAAISTKTAOTIKYGVAMGPQAAEVAAGVAK 60
 ; GENERAL INFORMATION:
 ; APPLICANT: Thonard, Joelle
 ; TITLE OF INVENTION: HAEMOPHILUS INFLUENZAE BASE202
 ; FILE REFERENCE: BM45419
 ; CURRENT APPLICATION NUMBER: US/10/380,817
 ; PRIOR FILING DATE: 2003-03-18
 ; PRIOR APPLICATION NUMBER: PCT/EP01/10979
 ; PRIOR FILING DATE: 2001-09-18
 ; PRIOR APPLICATION NUMBER: GB 0022992.2
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 273
 ; TYPE: PRF
 ; ORGANISM: No. US20040039169A1-typeable Haemophilus influenzae
 US-10-380-817-2
 Db 241 DNKSKAIQDFVKAAYQTDVEAEAKKQFQGVIKGW 276
 QY 241 DNKSKAIQDFVKAAYQTDVEAEAKKQFQGVIKGW 276
 Db 181 PKKLIVKEVDTSVAARAIDVDLAVVNNNTYAGVGLTASENGVEVEDKSPYNIIVARA 240
 QY 181 PKKLIVKEVDTSVAARAIDVDLAVVNNNTYAGVGLTASENGVEVEDKSPYNIIVARA 240
 Db 121 LAGSTKIKTLNELKDGTIAVPNDPSLALILLEKQGLIKKXNTNLFSTLTDIVEN 180
 QY 121 LAGSTKIKTLNELKDGTIAVPNDPSLALILLEKQGLIKKXNTNLFSTLTDIVEN 180
 Db 121 LAGSTKIKTLNELKDGTIAVPNDPSLALILLEKQGLIKKXNTNLFSTLTDIVEN 180
 QY 181 PKKLIVKEVDTSVAARAIDVDLAVVNNNTYAGVGLTASENGVEVEDKSPYNIIVARA 240
 Db 181 PKKLIVKEVDTSVAARAIDVDLAVVNNNTYAGVGLTASENGVEVEDKSPYNIIVARA 240

RESULT 2
 US-10-380-817-2
 ; Sequence 2, Application US/10380817
 ; Publication No. US20040039169A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Thonard, Joelle
 ; TITLE OF INVENTION: HAEMOPHILUS INFLUENZAE BASE202
 ; FILE REFERENCE: BM45419
 ; CURRENT APPLICATION NUMBER: US/10/380,817
 ; PRIOR FILING DATE: 2003-03-18
 ; PRIOR APPLICATION NUMBER: PCT/EP01/10979
 ; PRIOR FILING DATE: 2001-09-18
 ; PRIOR APPLICATION NUMBER: GB 0022992.2
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 273
 ; TYPE: PRF
 ; ORGANISM: No. US20040039169A1-typeable Haemophilus influenzae
 US-10-380-817-2

Query Match 72.8%; Score 1007.5; DB 12; Length 273;
 Best Local Similarity 71.7%; Pred. No. 1.4e-83;
 Matches 198; Conservative 27; Mismatches 48; Indels 3; Gaps 1;

QY 1 MNFGKINGICALASGIALAGCSNOSNEPAAISTKTAOTIKYGVAMGPQAAEVAAGVAK 60
 ; GENERAL INFORMATION:
 ; APPLICANT: Thonard, Joelle
 ; TITLE OF INVENTION: HAEMOPHILUS INFLUENZAE BASE202
 ; FILE REFERENCE: BM45419
 ; CURRENT APPLICATION NUMBER: US/10/380,817
 ; PRIOR FILING DATE: 2003-03-18
 ; PRIOR APPLICATION NUMBER: PCT/EP01/10979
 QY 121 LAGSTKIKTLNELKDGTIAVPNDPSLALILLEKQGLIKKXNTNLFSTLTDIVEN 180

Db 118 LAGSTKIKTLNELKDGTIAVPNDPSLALILLEKQGLIKKXNTNLFSTLTDIVEN 177
 QY 181 PKKLIVKEVDTSVAARAIDVDLAVVNNNTYAGVGLTASENGVEVEDKSPYNIIVARA 240
 Db 178 PKKLIVKEVDTSVAARAIDVDLAVVNNNTYAGVGLTASENGVEVEDKSPYNIIVARA 237
 QY 241 DNKSKAIQDFVKAAYQTDVEAEAKKQFQGVIKGW 276
 Db 238 DNKSKAIQDFVKAAYQTDVEAEAKKQFQGVIKGW 273

RESULT 3
 US-10-380-817-4
 ; Sequence 4, Application US/10380817
 ; Publication No. US20040039169A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Thonard, Joelle
 ; TITLE OF INVENTION: HAEMOPHILUS INFLUENZAE BASE202
 ; FILE REFERENCE: BM45419
 ; CURRENT APPLICATION NUMBER: US/10/380,817
 ; PRIOR FILING DATE: 2003-03-18
 ; PRIOR APPLICATION NUMBER: PCT/EP01/10979
 ; PRIOR FILING DATE: 2001-09-18
 ; PRIOR APPLICATION NUMBER: GB 0022992.2
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 273
 ; TYPE: PRF
 ; ORGANISM: No. US20040039169A1-typeable Haemophilus influenzae
 US-10-380-817-4

Query Match 72.8%; Score 1007.5; DB 12; Length 273;
 Best Local Similarity 71.7%; Pred. No. 1.4e-83;
 Matches 198; Conservative 27; Mismatches 48; Indels 3; Gaps 1;

QY 1 MNFGKINGICALASGIALAGCSNOSNEPAAISTKTAOTIKYGVAMGPQAAEVAAGVAK 60
 ; GENERAL INFORMATION:
 ; APPLICANT: Thonard, Joelle
 ; TITLE OF INVENTION: HAEMOPHILUS INFLUENZAE BASE202
 ; FILE REFERENCE: BM45419
 ; CURRENT APPLICATION NUMBER: US/10/380,817
 ; PRIOR FILING DATE: 2003-03-18
 ; PRIOR APPLICATION NUMBER: PCT/EP01/10979
 QY 121 LAGSTKIKTLNELKDGTIAVPNDPSLALILLEKQGLIKKXNTNLFSTLTDIVEN 180
 Db 118 LAGSTKIKTLNELKDGTIAVPNDPSLALILLEKQGLIKKXNTNLFSTLTDIVEN 177
 QY 181 PKKLIVKEVDTSVAARAIDVDLAVVNNNTYAGVGLTASENGVEVEDKSPYNIIVARA 240
 Db 178 PKKLIVKEVDTSVAARAIDVDLAVVNNNTYAGVGLTASENGVEVEDKSPYNIIVARA 237
 QY 241 DNKSKAIQDFVKAAYQTDVEAEAKKQFQGVIKGW 276
 Db 238 DNKSKAIQDFVKAAYQTDVEAEAKKQFQGVIKGW 273

RESULT 4
 US-10-380-817-8
 ; Sequence 8, Application US/10380817
 ; Publication No. US20040039169A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Thonard, Joelle
 ; TITLE OF INVENTION: HAEMOPHILUS INFLUENZAE BASE202
 ; FILE REFERENCE: BM45419
 ; CURRENT APPLICATION NUMBER: US/10/380,817
 ; PRIOR FILING DATE: 2003-03-18
 ; PRIOR APPLICATION NUMBER: PCT/EP01/10979

Tue Jun 22 11:04:04 2004

us-10-018-672-2.rapb

Page 3

```

; PRIOR APPLICATION NUMBER: GB 0022992.2
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 273
; TYPE: PRT
; ORGANISM: No. US20040039169A1-typeable Haemophilus influenzae
US-10-380-817-8

```

```

Query Match      72.8%; Score 1007.5; DB 12; Length 273;
Best Local Similarity 71.7%; Pred. No. 1.4e-83;
Matches 198; Conservative 27; Mismatches 48; Indels 3; Gaps 1;

```

```

QY 1 MNFGKINGICALASGIALAGCSNOSNEPAISTKTAOTIKYGVAGPEQAVAEVAGVAK 60
DB 1 MKLQOLFALTAIALSLVLTGCKEDKKPEAA--AAPLKIKGVWMSGPHQVAEIAKAK 57
QY 61 EKYNTLVELVEFNDYAMNSAVSKGELDANAMQHKPYLEKSOEKGNNLVYVNTFYYP 120
DB 58 EKYGLDVQFVEFNDYALPNEAVSKGDLNANMOKHPYLEDKAKANNLVYVNTFYYP 117
QY 121 LAGYSTKIKTLNELKDGATIAVPNDPSNLARALILIEKQGLIKKDNNTLSTLTDIYEN 180
DB 118 LAGYSKKIKNNVELDGAQAVVNDPTNRRALILIEKQGLIKKDNNTLSTLTDIYEN 177
QY 181 PKKVIKEVDTSVAARAIDVDLAVVNNNYAGVGGLTSENGVFVEDKDSFYNNITVARA 240
DB 178 PKKLNTEVDTSVAARALDDVDLAVVNNNYAGVGGLNADDDGVFVEDKDSFYNNITVRSRT 237
QY 241 DNKDSKAIDPFYKAYOTDEVEAEAKKQFQDGVYKGM 276
DB 238 DNKDSKAIVDPFYKSYOTEVEYQEAQKHFRDGVYKGM 273

```

```

RESULT 5
US-10-380-817-10
; Sequence 10, Application US/10380817
; Publication No. US20040039169A1
; GENERAL INFORMATION:
; APPLICANT: Thonard, Joelle
; TITLE OF INVENTION: HAEMOPHILUS INFLUENZAE BASR202
; FILE REFERENCE: BM45419
; CURRENT FILING DATE: 2003-03-18
; PRIOR FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: PCT/EP01/10979
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: GB 0022992.2
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 273
; TYPE: PRT
; ORGANISM: No. US20040039169A1-typeable Haemophilus influenzae
US-10-380-817-10

```

```

Query Match      72.8%; Score 1007.5; DB 12; Length 273;
Best Local Similarity 71.7%; Pred. No. 1.4e-83;
Matches 198; Conservative 27; Mismatches 48; Indels 3; Gaps 1;

```

```

QY 1 MNFGKINGICALASGIALAGCSNOSNEPAISTKTAOTIKYGVAGPEQAVAEVAGVAK 60
DB 1 MKLQOLFALTAIALSLVLTGCKEDKKPEAA--AAPLKIKGVWMSGPHQVAEIAKAK 57
QY 61 EKYNTLVELVEFNDYAMNSAVSKGELDANAMQHKPYLEKSOEKGNNLVYVNTFYYP 120
DB 58 EKYGLDVQFVEFNDYALPNEAVSKGDLNANMOKHPYLEDKAKANNLVYVNTFYYP 117
QY 121 LAGYSTKIKTLNELKDGATIAVPNDPSNLARALILIEKQGLIKKDNNTLSTLTDIYEN 180
DB 118 LAGYSKKIKNNVELDGAQAVVNDPTNRRALILIEKQGLIKKDNNTLSTLTDIYEN 177

```

```

QY 181 PKKVIKEVDTSVAARAIDVDLAVVNNNYAGVGGLTSENGVFVEDKDSFYNNITVARA 240
DB 178 PKKLNTEVDTSVAARALDDVDLAVVNNNYAGVGGLNADDDGVFVEDKDSFYNNITVRSRT 237
QY 241 DNKDSKAIDPFYKAYOTDEVEAEAKKQFQDGVYKGM 276
DB 238 DNKDSKAIVDPFYKSYOTEVEYQEAQKHFRDGVYKGM 273

```

```

RESULT 6
US-09-815-242-11067
; Sequence 11067, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Onisen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11067
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-815-242-11067

```

```

Query Match      72.8%; Score 1006.5; DB 9; Length 273;
Best Local Similarity 71.4%; Pred. No. 1.7e-83;
Matches 197; Conservative 28; Mismatches 48; Indels 3; Gaps 1;

```

```

QY 1 MNFGKINGICALASGIALAGCSNOSNEPAISTKTAOTIKYGVAGPEQAVAEVAGVAK 60
DB 1 MKLQOLFALTAIALSLVLTGCKEDKKPEAA--AAPLKIKGVWMSGPHQVAEIAKAK 57
QY 61 EKYNTLVELVEFNDYAMNSAVSKGELDANAMQHKPYLEKSOEKGNNLVYVNTFYYP 120
DB 58 EKYGLDVQFVEFNDYALPNEAVSKGDLNANMOKHPYLEDKAKANNLVYVNTFYYP 117
QY 121 LAGYSTKIKTLNELKDGATIAVPNDPSNLARALILIEKQGLIKKDNNTLSTLTDIYEN 180
DB 118 LAGYSKKIKNNVELDGAQAVVNDPTNRRALILIEKQGLIKKDNNTLSTLTDIYEN 177
QY 181 PKKVIKEVDTSVAARAIDVDLAVVNNNYAGVGGLTSENGVFVEDKDSFYNNITVARA 240
DB 178 PKKLNTEVDTSVAARALDDVDLAVVNNNYAGVGGLNADDDGVFVEDKDSFYNNITVRSRT 237
QY 241 DNKDSKAIDPFYKAYOTDEVEAEAKKQFQDGVYKGM 276
DB 238 DNKDSKAIVDPFYKSYOTEVEYQEAQKHFRDGVYKGM 273

```

```
RESULT 7
US-10-282-122A-58240
; Sequence 58240, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 58240
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-282-122A-58240

Query Match      72.8%; Score 1006.5; DB 12; Length 273;
Best Local Similarity 71.4%; Pred. No. 1.7e-83;
Matches 197; Conservative 28; Mismatches 48; Indels 3; Gaps 1;

QY 1 MNPFGKINGICALASGIALAGCSNQSNEPAISKTAQTIKVGWAGPEGAVAEVAAGQYAK 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1 MKLQKQFALTAIALAVLTGCKEKKPEAA--AAPLKIKVWMSGPEHQVAEIAKAK 57
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 EKXNLVLEVFENDYAMPNSAVSKGELDANAMQHPYLEKDSOEKGLNNLVVIGNTFYYP 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 58 EKXGLDVQVFENFDYALPNEAVSKGDLNANMOKHPYLEDKAKKLNINLVIGNTFYYP 117
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 LAGYSTKIKTLNELKDGATIAVPNDPSNLARALLILEKQGLIKLKNNTLSTTLDIYEN 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 118 LAGYSKKIKVNELODGAQVVPNDPTNRGRALLILEKQGLIKLKDANNLSTVLDIYEN 177
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 LAGYSTKIKTLNELKDGATIAVPNDPSNLARALLILEKQGLIKLKNNTLSTTLDIYEN 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 118 LAGYSKKIKVNELODGAQVVPNDPTNRGRALLILEKQGLIKLKDANNLSTVLDIYEN 177
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 181 PKKLVIKEVDTSVAARAIDVDVLAVNNNYAGVGLTASNGVFVEDKSPYNIIVARA 240
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 178 PKKLNITEVDTSVAARAIDVDVLAVNNNYAGVGLTASNGVFVEDKSPYNIIVARS 237
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 241 DNKDSKAIDPFVKAAYQTEVEAEAKKQFKDGVYKGM 276
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
DB 238 DNKDSKAVIDPFKSYQTEEVYQEAQKHFQDGVYKGM 273
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
RESULT 8
US-10-380-817-6
; Sequence 6, Application US/10380817
; Publication No. US20040039169A1
; GENERAL INFORMATION:
; APPLICANT: Thonnard, Joelle
; TITLE OF INVENTION: HAEMOPHILUS INFLUENZAE BASB202
; FILE REFERENCE: BM45419
; CURRENT APPLICATION NUMBER: US/10/380,817
; PRIOR FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: PCT/EP01/10979
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: GB 0022992.2
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatsSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 273
; TYPE: PRT
; ORGANISM: No. US20040039169A1-typeable Haemophilus influenzae
US-10-380-817-6

Query Match      72.8%; Score 1006.5; DB 12; Length 273;
Best Local Similarity 71.7%; Pred. No. 1.7e-83;
Matches 198; Conservative 26; Mismatches 49; Indels 3; Gaps 1;

QY 1 MNPFGKINGICALASGIALAGCSNQSNEPAISKTAQTIKVGWAGPEGAVAEVAAGQYAK 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1 MKLQKQFALTAIALAVLTGCKEKKPEAA--AAPLKIKVWMSGPEHQVAEIAKAK 57
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 EKXNLVLEVFENDYAMPNSAVSKGELDANAMQHPYLEKDSOEKGLNNLVVIGNTFYYP 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 58 EKXGLDVQVFENFDYALPNEAVSKGDLNANMOKHPYLEDKAKKLNINLVIGNTFYYP 117
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 LAGYSTKIKTLNELKDGATIAVPNDPSNLARALLILEKQGLIKLKNNTLSTTLDIYEN 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 118 LAGYSKKIKVNELODGAQVVPNDPTNRGRALLILEKQGLIKLKDANNLSTVLDIYEN 177
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 181 PKKLVIKEVDTSVAARAIDVDVLAVNNNYAGVGLTASNGVFVEDKSPYNIIVARA 240
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 178 PKKLNITEVDTSVAARAIDVDVLAVNNNYAGVGLTASNGVFVEDKSPYNIIVARS 237
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 241 DNKDSKAIDPFVKAAYQTEVEAEAKKQFKDGVYKGM 276
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 238 DNKDSKAVIDPFKSYQTEEVYQEAQKHFQDGVYKGM 273
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 9
US-10-282-122A-67387
; Sequence 67387, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
```

PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 67387
LENGTH: 276
TYPE: PRT
ORGANISM: Pasteurella multocida
US-10-282-122A-67387

Query Match 64.2% Score 888; DB 12; Length 276;
Best Local Similarity 60.2%; Pred. No. 1,1e-72;
Matches 168; Conservative 55; Mismatches 50; Indels 6; Gaps 3;
QY 1 MFMGKINGICATLALSGIALAGCSNQSNEPAIS---KTAQTIVGVMAQPEQVAEVAQ 57
Db 1 MKRTKLFGLATVLSAVALAGC--KDKPAAAPQAPPAKRLTVGWTGAQVTEVAAK 58
QY 58 VAEKKNLYVEVFNDFYAMPNSAVSKGELDANAMQHKPYLEKDSOEKGLNLVIGNTF 117
Db 59 IAEKKNIDVKEVLEFVEYQPNDAITKGDIDANAFQHKRYMDKEVQRYG-KLAIYGNF 117
QY 118 VYELAGVSTKIKTLNLEKQATIAVNDPSNLARALILEKQGLIKLKNNTLFFSTLTDI 177
Db 118 VFPIAAVSKIKIVSELQDQATVAVPNNPSNLGRALLLEKQGLIKLKPNSLFSISIV 177
QY 178 VENPKLVIEVDTSVAAPALDDVDLAVNNNTAGVGLTASENGVEVEDKDSPPYNIIV 237
Db 178 IENPKLVIEVDSGLPRLDDVDPAITNNNTAVQGLTAEKDGIFVEDKDSPPYNIIV 237
QY 238 ARADNDSKALIDFVKAYQTDVEAEAKQFKDGVIKGM 276
Db 238 SREDNKNDAIKDFVAFQTEEVYQDALHFGQGVKGM 276

RESULT 10
US-10-282-122A-76195
Sequence 76195, Application US//10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 76195
LENGTH: 271
TYPE: PRT
ORGANISM: Salmonella typhi
US-10-282-122A-76195

Query Match 57.4% Score 793.5; DB 12; Length 271;
Best Local Similarity 56.5%; Pred. No. 4.7e-64;
Matches 156; Conservative 44; Mismatches 65; Indels 11; Gaps 3;
QY 3 FKGKINGICATLALSGIALAGCSNQSNEPAISKTAQTIVGVMAQPEQVAEVAQVKEK 62
Db 5 FKTPAAGALIGSIALAGCGGDEKDP-----NHRKGVTVAEQVQVAEQVAKVKEK 56
QY 63 YNLTEVLEFVFNDFYAMPNSAVSKGELDANAMQHKPYLEKDSOEKGLNLVIGNTFVYPLA 122
Db 57 YGLDVEVLETFNDYVLPENALSKGPDIDANAFQHKRYLDDQIDRQY-KLVSVGKTFVYPLA 115
QY 123 GYSTIKITLNLKQATIAVNDPSNLARALILEKQGLIKLKNNTLFFSTLTDIENPK 182
Db 116 GYSKIKISLDELKQGSQVAVNPDPNTNIGRSLLIQKGLIKLKGVSLLPSTLSDIENPK 175
QY 183 KLVIKEVDTSVAAPALDDVDLAVNNNTAGVGLTASENGVEVEDKDSPPYNIIVARA 240
Db 176 NKVIELEAPQLPFSLDDAQALAVINTTASQGLTPAKDGIVEDKDSPPYNIIVTRE 235
QY 241 DNKDSKALIDFVKAYQTDVEAEAKQFKDGVIKGM 276
Db 236 DNKDAENVKKEFVQAYQSDVEVEAANKVNGGAVKGM 271

RESULT 11
US-09-815-242-10050
Sequence 10050, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242

```
/ CURRENT FILING DATE: 2001-03-21
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/257,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 14110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 10050
/ LENGTH: 271
/ TYPE: PRF
/ ORGANISM: Escherichia coli
US-09-815-242-10050
```

```
Query Match 56.6%; Score 782.5; DB 9; Length 271;
Best Local Similarity 55.1%; Pred. No. 4,7e-63;
Matches 152; Conservative 46; Mismatches 67; Indels 11; Gaps 3;
```

```
QY 3 FGIKNGICALASGIALAGCSNQSNEPAAISKTAQITIKVGMGPBGQVAEVAQVAKKK 62
DB 5 FKTFAAGALIGSLALVGCQDEKDP-----NHIKGVIVGAEQVAEVAQVAKKK 56
QY YNLTFELVEFNDYAMPNSAVSKGELDANAMQHKRYLEKDSQEKGLNNLVIVGNTFVYPLA 122
DB 57 YGIDVELVTFNDYLPHEALSKGIDIDANAFQHKPYLDQQLKRGY-KLVAVGNTFVYPLA 115
QY 123 GYSTIKTLNELKQATITAVPNDPSNLARALLLEKQGLIKLQNTNLFSTTLDIYENK 182
DB 116 GYSKIKSLDELQSGQVAVPNDPTNLGRSLILQKVGILKLDQVGLPTVLDVVENPK 175
QY 183 KLVIKEDVTSVAARAID--VDLAVVNNNYAGQVGLTASENGVEDEKSPYNIIVARA 240
DB 176 NLKIVLEAPQLPSLDDAQIALAVINTTYSQIGLTPACDGLFVEDKESPYNLIVTRE 235
QY 241 DNKSKAIQDFVKAQYQDEVEAEAKKQFKQGVIKGW 276
DB 236 DNKDAENVKKEFQAYQSDVEYEAANKVFNGAVKGW 271
```

```
RESULT 12
US-10-282-122A-43291
/ Sequence 43291, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: EUTRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
```

```
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 43291
/ LENGTH: 271
/ TYPE: PRF
/ ORGANISM: Escherichia coli
US-10-282-122A-43291
```

```
Query Match 56.6%; Score 782.5; DB 12; Length 271;
Best Local Similarity 55.1%; Pred. No. 4,7e-63;
Matches 152; Conservative 46; Mismatches 67; Indels 11; Gaps 3;
```

```
QY 3 FGIKNGICALASGIALAGCSNQSNEPAAISKTAQITIKVGMGPBGQVAEVAQVAKKK 62
DB 5 FKTFAAGALIGSLALVGCQDEKDP-----NHIKGVIVGAEQVAEVAQVAKKK 56
QY YNLTFELVEFNDYAMPNSAVSKGELDANAMQHKRYLEKDSQEKGLNNLVIVGNTFVYPLA 122
DB 57 YGIDVELVTFNDYLPHEALSKGIDIDANAFQHKPYLDQQLKRGY-KLVAVGNTFVYPLA 115
QY 123 GYSTIKTLNELKQATITAVPNDPSNLARALLLEKQGLIKLQNTNLFSTTLDIYENK 182
DB 116 GYSKIKSLDELQSGQVAVPNDPTNLGRSLILQKVGILKLDQVGLPTVLDVVENPK 175
QY 183 KLVIKEDVTSVAARAID--VDLAVVNNNYAGQVGLTASENGVEDEKSPYNIIVARA 240
DB 176 NLKIVLEAPQLPSLDDAQIALAVINTTYSQIGLTPACDGLFVEDKESPYNLIVTRE 235
QY 241 DNKSKAIQDFVKAQYQDEVEAEAKKQFKQGVIKGW 276
DB 236 DNKDAENVKKEFQAYQSDVEYEAANKVFNGAVKGW 271
```

```
RESULT 13
US-10-282-122A-55826
/ Sequence 55826, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: EUTRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
```

PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 55826
LENGTH: 271
TYPE: PR1
ORGANISM: Enterobacter cloacae
US-10-282-122A-55826

Query Match 56.4%; Score 780.5; DB 12; Length 271;
Best Local Similarity 56.7%; Pred. No. 7.2e-63;
Matches 152; Conservative 43; Mismatches 62; Indels 11; Gaps 3;
QY 11 ALASGIALGCSNQSNEPAISKTAAQTITKVGMAPEQVAEVAQVAKKNTLVEIV 70
DB 13 ALGSLALVGCGGDEKDP-----NHKVGIVGAEOQVAEVAQVAKKNTLVEIV 64
QY 71 EFNDAVMPNSAVSKGELDANAMQHKPYLEKDSQEKGLNNLVITVGNTPFYVPLAGSYTKIT 130
DB 65 TENDVYLPNEALSKGIDANAFQHKPYLDQCIKDRGY-KLVVGNTPFYVPIAGYSKTKIS 123
QY 131 LNLKGGATIAVNDPSNLARALILIEKQGLIKLKNNTLFTTLDIVNPKKVIKEVD 190
DB 124 LDELQGSQVAAYNDPTNIGRSLLILQKVGILKLDGVGLLPVLDVTENPKMLKIVELE 183
QY 191 TSVARAID--VDLAVNNVAGQVGLTASENGVEVEDKDSPYNIIYARADNDSKAI 248
DB 184 APQLPSLDDAQALAVITTYASQIGLTPAKDGIFFVEDKDSPYNIIYARADNDSKAI 243
QY 249 QDFVKAQYQDVEVAEAKKQFKDGVTKGW 276
DB 244 KKEVQAYQSDDEVYQEAANKVFNGAVKGW 271

RESULT 14
US-10-282-122A-59600
Sequence 59600, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 59600
LENGTH: 271
TYPE: PR1
ORGANISM: Klebsiella pneumoniae
US-10-282-122A-59600

Query Match 56.1%; Score 775.5; DB 12; Length 271;
Best Local Similarity 55.2%; Pred. No. 2.1e-62;
Matches 153; Conservative 45; Mismatches 68; Indels 11; Gaps 3;
QY 2 NFGKNGICALASGIALGCSNQSNEPAISKTAAQTITKVGMAPEQVAEVAQVAKK 61
DB 4 NFKPRAVNGALIGSLALVGCGGDEKDP-----NHKVGIVGAEOQVAEVAQVAKK 55
QY 62 KYNLTVEIVEFNDAVMPNSAVSKGELDANAMQHKPYLEKDSQEKGLNNLVITVGNTPFYVPL 121
DB 56 KYGLDVEIVTENDVYLPNEALSKGIDANAFQHKPYLDQCIKDRGY-KLVVGNTPFYVPI 114
QY 122 AGYSTKIKTNLDELQGSQVAAYNDPTNIGRSLLILQKVGILKLDGVGLLPVLDVTENPKMLKIVELE 181
DB 115 AGYSKTKISLDELQGSQVAAYNDPTNIGRSLLILQKVGILKLDGVGLLPVLDVTENPKMLKIVELE 174
QY 182 KCLVKEVDTSVARAID--VDLAVNNVAGQVGLTASENGVEVEDKDSPYNIIYARADNDSKAI 239
DB 175 KNLKIVELEAPQLRSLDDAQALAVITTYASQIGLTPAKDGIFFVEDKDSPYNIIYARADNDSKAI 234
QY 240 ADNDSKAIQDFVKAQYQDVEVAEAKKQFKDGVTKGW 276
DB 235 EDNKDAENVKQFVQAYQSDDEVYQEAANKVFNGAVKGW 271

RESULT 15
US-10-282-122A-68391
Sequence 68391, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A

```

; CURRENT APPLICATION NUMBER: US/10/282,122A
; FILE REFERENCE: EUTRA.034A
; CURRENT APPLICATION NUMBER: 2003-02-20
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 68391
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Proteus mirabilis
; US-10-282-122A-68391

Query Match      54.8%; Score 757.5; DB 12; Length 271;
Best Local Similarity 52.7%; Pred. No. 9.1e-61;
Matches 147; Conservative 49; Mismatches 70; Indels 13; Gaps 4;

QY      1 MNEFGKINGICLALAGCSNOSNEPAISKTAQTIKVGMAGPEQVAEVAQVAK 60
DB      3 LKFKSLAVASALVGLALAGCGEKXP-----NHIRGVISGSEQVAEVAQVAK 54
QY      61 EKXNLTELVEFNDYAMPNSAVSKGELDANAMQHKREYLEKDSQEGKLN-LVIVGNTFPY 119
DB      55 DKYGLDELVTYFNDYVLPNSALSKGDIIDNAFQHKRPL--DQIKRNTKITVAGTFFY 112
QY      120 PLAGYSTIKITLNLKOGATIAVNDPSNLAARALILLEKQGLIKLXONTNLFSITLDIYE 179
DB      113 PLAGYSKITITLADLPDAQVAIIPNDPTNLGRSILLLEKQGLVLEKGVGLLPTFKLDIIE 172
QY      180 NPKKLVKEVDTSVAARAID--VDLAVNNNNVAGVGLTASENGVEVEDKDSPYNITIV 237
DB      173 NPKKLVLEAPQVLPNSLDQKITYLAVINTTASQVNLTPADGIFVEDKDSPYNITIV 232
QY      238 ARADKSKAIQDFYKAYQTDVEVAEAKQFKQGVIXGM 276
DB      233 ARDNKQDENYKKEFIQSQYOTDEVDSAAKIFNGGAVXGM 271

RESULT 16
US-10-282-122A-78174
; Sequence 78174, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
```

```

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EUTRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT APPLICATION NUMBER: 2003-02-20
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 78174
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Yersinia pestis
; US-10-282-122A-78174

Query Match      54.8%; Score 757.5; DB 12; Length 271;
Best Local Similarity 53.2%; Pred. No. 9.1e-61;
Matches 148; Conservative 48; Mismatches 71; Indels 11; Gaps 3;

QY      1 MNEFGKINGICLALAGCSNOSNEPAISKTAQTIKVGMAGPEQVAEVAQVAK 60
DB      3 LKFKSLAIAISALICTITLVGCGPTEKAP-----NHIKGVTVGAEQVAEVAQVAK 54
QY      61 EKXNLTELVEFNDYAMPNSAVSKGELDANAMQHKREYLEKDSQEGKLN-LVIVGNTFPY 120
DB      55 EKYGLDELVTYFNDYVLPNSALSKGDIIDNAFQHKRPL--DQIKRNTKITVAGTFFY 113
QY      121 LAGYSTIKITLNLKOGATIAVNDPSNLAARALILLEKQGLIKLXONTNLFSITLDIYE 180
DB      114 IAGYSKITISLDELQPSQVALLPNDPTNLGRSILLLOSQVGLIKLXONTNLFSITLDIYE 173
QY      181 PKKLVKEVDTSVAARAID--VDLAVNNNNVAGVGLTASENGVEVEDKDSPYNITIV 238
DB      174 PKKLVLEAPQVLPNSLDQKITYLAVINTTASQVNLTPADGIFVEDKDSPYNITIV 233
QY      239 RADKSKAIQDFYKAYQTDVEVAEAKQFKQGVIXGM 276
DB      234 REDKQDENYKKEFIQSQYOTDEVDSAAKIFNGGAVXGM 271

RESULT 17
US-10-282-122A-77174
; Sequence 77174, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
```


PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: Patent version 3.1
SEQ ID NO: 47704
LENGTH: 295
TYPE: PRT
ORGANISM: Burkholderia cepacia
US-10-282-122A-47704

Query Match 40.6%; Score 562; DB 12; Length 295;
Best Local Similarity 44.5%; Pred. No. 7.1e-43;
Matches 118; Conservative 47; Mismatches 88; Indels 12; Gaps 3;

QY 12 LASGIALAGCSNOSNEPAISKTAAQTIVGWAGPEQVAEVAQVAKKYNLTVEIVE 71
DB 37 LAGALTLTSGAHAD-----KVIKGVTAAGPDSVWQVAKKKEGKLNIVKIE 88
QY 72 FNDYAMPNSAVSKGELDANAMQHKPYLEKDSOEKGLNLYVGNTFVYPLAGYSTKIKTL 131
DB 89 FNYVFPNALDSDGIDANSFOHPYLDQVOKRGY-KIVSAGLYTSPICVYSKRPKSL 147
QY 132 NEUKDGTIAVNPDPENLARALILEKQGLIKLKONTNL---FSTTLDIVENPKLVKE 188
DB 148 KDLPOGKXAVPDPENRALLILQTOGVITLKAAGTGGNNATVLIAINPKKIKISE 207
QY 189 VDSVAAARADVDVLAANNVNTYAGVGLTASENGVPVEDKSPYNTIIVAAADNKKSKAI 248
DB 208 LDAADPRLVSDVAIVNTYALANLOFTKDALALSLTSPYANLIIAIAKODOPW 267
QY 249 QDEVKAYQTEDEVEAEAKKQPKDGI 273
DB 268 KKLIVKAYQSPVEVEKFKQPKGSMV 292

RESULT 20

US-10-282-122A-51371
Sequence 51371, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: Patent version 3.1
SEQ ID NO: 51371
LENGTH: 262
TYPE: PRT
ORGANISM: Bordetella pertussis
US-10-282-122A-51371

Query Match 39.8%; Score 551; DB 12; Length 262;
Best Local Similarity 43.1%; Pred. No. 6.1e-42;
Matches 118; Conservative 52; Mismatches 86; Indels 18; Gaps 5;

QY 1 NMPGKINGICALASGIALAGCSNOSNEPAISKTAAQTIVGWAGPEQVAEVA-GOVA 59
DB 3 MNPVR---SALLASFLAGAAQ-----AKLVGNTQVPHAILLEVPALAA 48
QY 60 KKKYNLTVEIVEFNDYAMPNSAVSKGELDANAMQHKPYLEKDSOEKGLNLYVGNTFVY 119
DB 49 KEGVLDIKV--FTDYVQPNQLADKQIDANFFQHPYLDTPFKDR-KTNLVSVGVVHVE 105
QY 120 PLAGYSTKIKTLNELKQDATTAVNPDPENLARALILEKQGLIKLKONTNLFTTLDIVE 179
DB 106 PGGSKIKIKSLAEKQDATTIIPNDPSNGRALLILQKGLKADPSIIVATPIDINE 165
QY 180 NPKLVIVKEDTSSVAAARADVDVLAANNVNTYAGVGLTASENGVPVEDKSPYNTIIVAR 239
DB 166 NPKKLFRELEAAMLPSEFDDLDLALINTYALEAGVPTPRDALFIEGADSPYANLVAAR 225
QY 240 ADNKKSKAIQDFKAYQTEDEVEAEAKKQPKDGI 273
DB 226 PDNKAAPVAKLVNHLHSEAVRKRTIRKYGAVV 259

RESULT 21

US-10-282-122A-50375
Sequence 50375, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20

```
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 50375
/ LENGTH: 272
/ TYPE: PRT
/ ORGANISM: Burkholderia mallei
/ US-10-282-122A-50375
```

```
Query Match      39.7%; Score 549.5; DB 12; Length 272;
Best Local Similarity 42.5%; Pred. No. 8,8e-42;
Matches 114; Conservative 52; Mismatches 91; Indels 11; Gaps 3;

QY      9 ICALSGIALAGCSNOSNEPAISKTAAQTIKVGMAGPEQAVAEVAGVAKKXNLTVE 68
Db      10 VSAVAGAAALSVS-----VGAQADPKVIVAGVDAQVQVQVAKKXGGLDVK 62

QY      69 LVEFNDYAMPNSAVSGEGLDANAMQKYLEDSQSEKLNLYVGNFVYPLAGYSKI 128
Db      63 VIEFNDYVQPNALDSGDDANSFQHPYLSQVQRQRY-KIVSAGLYTISPIGYSKIF 121

QY      129 KTLMEHKDQATTAVPNDPSNLTARALILEKQGLIKLKXNTNL---FSTLTDIVENPKLV 185
Db      122 KSLKELPAKAKVALPNDPENENRALLILQGVILKKGAGGGSNAITVDAENPKKXK 181

QY      186 IREVTISVAARAIDVDTLAVVNNNYAGVGLTASENGVFVEDKSPYNNIIVARAUNKDS 245
Db      182 LTELDAQAQPRVLSIDVDAVINTNYALANLOPTKDALSLSTSPYNNLIIVAKKQDQ 241

QY      246 KAIQDFKAYQTDVEVAEAKKQFKDQVI 273
Db      242 PWTCKLVKAYQSEPEVKEFTKQFKGSMV 269
```

```
RESULT 22
US-10-282-122A-61127
/ Sequence 61127; Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Karl
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
```

```
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 61127
/ LENGTH: 256
/ TYPE: PRT
/ ORGANISM: Legionella pneumophila
/ FEATURE:
/ NAME/KEY: MISC FEATURE
/ LOCATION: (227)..(227)
/ OTHER INFORMATION: X=any amino acid
/ FEATURE:
/ NAME/KEY: MISC FEATURE
/ LOCATION: (241)..(241)
/ OTHER INFORMATION: X=any amino acid
/ FEATURE:
/ NAME/KEY: MISC FEATURE
/ LOCATION: (244)..(244)
/ OTHER INFORMATION: X=any amino acid
/ NAME/KEY: MISC FEATURE
/ LOCATION: (248)..(248)
/ OTHER INFORMATION: X=any amino acid
/ US-10-282-122A-61127
```

```
Query Match      39.4%; Score 545; DB 12; Length 256;
Best Local Similarity 45.7%; Pred. No. 2,1e-41;
Matches 118; Conservative 37; Mismatches 91; Indels 12; Gaps 3;

QY      16 ILAGCSNOSNEPAISKTAAQTIKVGMAGPEQAVAEVAGVAKKXNLTVEVENDY 75
Db      11 ISLVACS-----SKPANTLVIGTIGAPFETHELITPAQVAKKXNLTVEVENDY 61

QY      76 AMPNSAVSGEGLDANAMQKYLEDSQSEKLNLYVGNFVYPLAGYSKIKTLNELK 135
Db      62 NLPNEALDQGSIDAVNYOHLPYLKAALISHGY-DLOAIGRTFVYPMGYSKKYTLSELFP 120

QY      136 DQATTAVPNDPSNLTARALILEKQGLIKLKXNTNLFSTLTDIVENPKLVKEVDTSVAA 195
Db      121 ENGIIAVPNDPENENRALLILEKXALITLKNTTN--SGIQIBSNPKQFKKELDAQLP 178

QY      196 RAIDVDLAVVNNNYAGVGLTASENGVFVEDKSPYNNIIVARAUNKDSALQDFKAY 255
Db      179 RVLPDVDAVINTTALPAGLSPSKDALFTGKQSPYNNIIVIRBDTEXRPQLELFYKAL 238

QY      256 QTDVEVAEAKKQFKDQVI 273
Db      239 NSXEVKEXKXNLFGEDAI 256
```

```
RESULT 23
US-10-282-122A-61127
/ Sequence 61127; Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Karl
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
```


Db 175 ESAMLPRLVDVLDLMDINTNVALEAGLNPAKALVLEGASPYVNLVAPDNKDSIAQ 234
 QY 250 DFVAYQOTDEVEAEAKKQFDGVI 273
 Db 235 KLAAGLTSPVKAFAKKGAVL 258

RESULT 27

US-09-815-242-12110
 ; Sequence 12110, Application US/09815242
 ; Patent No. US20020061569A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Karl L.
 ; APPLICANT: Zysek, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John D.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 ; TITLE OF INVENTION: Identification of Essential Genes in
 ; TITLE OF INVENTION: Prokaryotes
 ; FILE REFERENCE: ELITRA.011A
 ; CURRENT APPLICATION NUMBER: US/09/815,242
 ; CURRENT FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; NUMBER OF SEQ ID NOS: 14110
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 12110
 ; LENGTH: 260
 ; TYPE: PRN
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-815-242-12110

Query Match 36.3%; Score 502.5; DB 9; Length 260;
 Best Local Similarity 42.6%; Pred. No. 1.6e-37;
 Matches 106; Conservative 52; Mismatches 84; Indels 7; Gaps 3;
 QY 29 AAIKSTAAQTIKYGWAGPEQAVAEVAGV---AKEKYNLTVELVEFNDYAMPNSAVSK 84
 Db 12 AALGITAQAASLUTVATPVPBHEILNVKPLAKGVLDKIK--FTDYYQPNVQVSE 69
 QY 85 GELDANAMOKPYLEKDSOEKGLNNLVIGNTFVYPLAGYSTKIKTLNELKOGATTAVPN 144
 Db 70 KRLDANFQHPYLDENFKAG-TDLVAVTVGVHIEPLGAYSSRYKLDLPSGATVVIEN 128
 QY 145 DPSNLARALITLEKOGILKIKONTNLFSTLTDIVENPKLVKEVDTSVAAARLIDVDLA 204
 Db 129 DATNGGRLLILLDKAGVITKDKNSITATPKDIVNPKIKIRLEAATLPRVLTVQDMA 188
 QY 205 VVNNNTAGQVGLTASENGVEFEDKSPYVNIIVARADNKSQAIODPVKAYQOTDEVEAEA 264
 Db 189 LINTNVALEAKLNPDKALALIEGSDSPYVNIIVARPDNDSQAMOKLAKALHSAEIKOFI 248
 QY 265 KQKQFDGVI 273
 Db 249 QEKYKAVV 257

RESULT 28

US-10-282-122A-66844
 ; Sequence 66844, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Karl
 ; APPLICANT: Zysek, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Foreyth, R.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78514
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 66844
 ; LENGTH: 260
 ; TYPE: PRN
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-10-282-122A-66844

Query Match 36.3%; Score 502.5; DB 12; Length 260;
 Best Local Similarity 42.6%; Pred. No. 1.6e-37;
 Matches 106; Conservative 52; Mismatches 84; Indels 7; Gaps 3;
 QY 29 AAIKSTAAQTIKYGWAGPEQAVAEVAGV---AKEKYNLTVELVEFNDYAMPNSAVSK 84
 Db 12 AALGITAQAASLUTVATPVPBHEILNVKPLAKGVLDKIK--FTDYYQPNVQVSE 69
 QY 85 GELDANAMOKPYLEKDSOEKGLNNLVIGNTFVYPLAGYSTKIKTLNELKOGATTAVPN 144
 Db 70 KRLDANFQHPYLDENFKAG-TDLVAVTVGVHIEPLGAYSSRYKLDLPSGATVVIEN 128
 QY 145 DPSNLARALITLEKOGILKIKONTNLFSTLTDIVENPKLVKEVDTSVAAARLIDVDLA 204
 Db 129 DATNGGRLLILLDKAGVITKDKNSITATPKDIVNPKIKIRLEAATLPRVLTVQDMA 188
 QY 205 VVNNNTAGQVGLTASENGVEFEDKSPYVNIIVARADNKSQAIODPVKAYQOTDEVEAEA 264
 Db 189 LINTNVALEAKLNPDKALALIEGSDSPYVNIIVARPDNDSQAMOKLAKALHSAEIKOFI 248
 QY 265 KQKQFDGVI 273
 Db 249 QEKYKAVV 257

RESULT 29
US-09-815-242-5146
Sequence 5146, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA, 011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5146
LENGTH: 259
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-815-242-5146

Query Match 36.0%; Score 497.5; DB 9; Length 259;
Best Local Similarity 41.4%; Pred. No. 4,5e-37;
Matches 110; Conservative 51; Mismatches 88; Indels 17; Gaps 4;
QY 9 ICALSGIALAGCSNQSNEPAISKTAAOTIKVGMAGPEQAVAE-VAGOVAKKKNLTV 67
DB 7 LTLALAVT-----SVSASNAEKLIIVATPIPHAEILILKPTLAKGVLDI 53
QY 68 ELVEFNDYAMPNSAVSKGELDANAMOHKPYLEKDSQEGKLNLLVIGNTFVYPPLAGYSTK 127
DB 54 KV--FTDYQVQVAVAEKRLDANVFQTLPYLENFNKGGTGLVTVVG-VHVEPFGYSRK 110
QY 128 IKTINELKOGATIAVNDPSNLARALILKEKGLIKLKDNTNLFSTTLDIVENPKKVIK 187
DB 111 YKSLALPBGATVAVIPNPSNLARALILKEKGLIKLKDNTNLFSTTLDIVENPKKVIK 170
QY 188 EVDTSVAAARAIIDVDLAVVNNNYAGOVGLTASENGVFEEDKDSPPYVNIIVARADNKD SKA 247
DB 171 ELESALLPRLVDQVLDLINTVYALAEAKLNPAKDALVLEDRDSPYVNYVAVARPNDKSDA 230
QY 248 IODPVKAYOTDEVEAEAKKQKFDGVI 273
DB 231 LKLSAALTSPVKAIFIEKTYAGAVV 256

RESULT 30
US-10-282-122A-43469
Sequence 43469, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos

APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA, 034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See file wrapper or PAM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: Patent in version 3.1
SEQ ID NO 43469
LENGTH: 259
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-10-282-122A-43469

Query Match 36.0%; Score 497.5; DB 12; Length 259;
Best Local Similarity 41.4%; Pred. No. 4,5e-37;
Matches 110; Conservative 51; Mismatches 88; Indels 17; Gaps 4;
QY 9 ICALSGIALAGCSNQSNEPAISKTAAOTIKVGMAGPEQAVAE-VAGOVAKKKNLTV 67
DB 7 LTLALAVT-----SVSASNAEKLIIVATPIPHAEILILKPTLAKGVLDI 53
QY 68 ELVEFNDYAMPNSAVSKGELDANAMOHKPYLEKDSQEGKLNLLVIGNTFVYPPLAGYSTK 127
DB 54 KV--FTDYQVQVAVAEKRLDANVFQTLPYLENFNKGGTGLVTVVG-VHVEPFGYSRK 110
QY 128 IKTINELKOGATIAVNDPSNLARALILKEKGLIKLKDNTNLFSTTLDIVENPKKVIK 187
DB 111 YKSLALPBGATVAVIPNPSNLARALILKEKGLIKLKDNTNLFSTTLDIVENPKKVIK 170
QY 188 EVDTSVAAARAIIDVDLAVVNNNYAGOVGLTASENGVFEEDKDSPPYVNIIVARADNKD SKA 247
DB 171 ELESALLPRLVDQVLDLINTVYALAEAKLNPAKDALVLEDRDSPYVNYVAVARPNDKSDA 230
QY 248 IODPVKAYOTDEVEAEAKKQKFDGVI 273
DB 231 LKLSAALTSPVKAIFIEKTYAGAVV 256

RESULT 31
US-10-282-122A-52976
Sequence 52976, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos

APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: EITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 52976
LENGTH: 263
TYPE: PRT
ORGANISM: Clostridium difficile
US-10-282-122A-52976
Query Match 35.7%; Score 494; DB 12; Length 263;
Best Local Similarity 41.7%; Pred. No. 9.7e-37;
Matches 111; Conservative 50; Mismatches 87; Indels 18; Gaps 6;
QY 11 ALASGIALA--GCSNCSNEPAISKTAAQTIRKVGWAGPEQVAEVAQVAKER-YNLTV 67
DB 10 ALVAIAISAVGCSNKED-----KKILVGAASNPHAKILEVAKPLKEKEGYDLEV 59
QY 68 ELVEENDYAMPNSAVKSGELDANAMQHKPYLEKDSQKGLNLUYVGNTPFYVPLAGYSTK 127
DB 60 KI--FDYVLPNPLALDEGSLDANFQHIPLLEBYEKGY-KLTYTSKVHIEPMGFISEK 116
QY 128 IKTLNELKDGATTAVPNDPSNLARALLILEKQGLIKLKNNTNLFSTTLDIVENPKLVYK 187
DB 117 VKSIDELKDGAVIAVPNDATNGARALKLAKKLLEVKDGLI--TKXDITNPNRNVYIK 174
QY 188 EVDTSVAARADIDVDLAIVNNVNAAGVGLTASENGVFVEDKSPYVNIIVARADKDSKA 247
DB 175 EWMNEQLPTVLKQVDGAVINSNVALTANLNPTKDAIVLESSPSYVNIILACRENNKDSK 234
QY 248 IODFVAXYQTDVEAEAKQKQFQDGV 273
DB 235 IKALSEAMNSKEVKEFIQDEYKGSIV 260

RESULT 32
US-10-282-122A-67627
Sequence 67627, Application US/10282122A

Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: EITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 67627
LENGTH: 256
TYPE: PRT
ORGANISM: Pseudomonas putida
US-10-282-122A-67627
Query Match 35.3%; Score 488; DB 12; Length 256;
Best Local Similarity 42.3%; Pred. No. 3.3e-36;
Matches 105; Conservative 48; Mismatches 89; Indels 6; Gaps 3;
QY 23 AISTKTAQTIRKVGWAGP--EQAVAEVAQVAKERKNTLVLEVFENDYAMPNSAVSKG 85
DB 9 AAVAAFSAQADLTVAATPVPFAELINFEKQLAKEGVELKVK--EFTDYIQPNVQVAK 66
QY 86 ELDANAMQHKPYLEKDSQKGLNLUYVGNTPFYVPLAGYSTKTKTLNELKDGATTAVPND 145
DB 67 RLDANFPQHPYLDDEFNFAKG-TSLVSVAGVHIEPLGVSTIKIKLDELSGATVIEND 125
QY 146 PNLARALLILEKQGLIKLKNNTNLFSTTLDIVENPKLVYKEVDSVAARAIDVDVLAV 205
DB 126 ANNGRALLILDDKAGVILKDKNKNILSTVKQVAENPNVAVFRELEAATTIRVITOVDAAL 185
QY 206 UNNNTAGVGLTASENGVFVEDKSPYVNIIVARADKDSKALODFVAXYQTDVEAEAK 265
DB 186 INTNVALEKLANPEKDALIEGSDSPYVNIIVARADKDSKLAALHSEFVKOFIT 245
QY 266 KQFQDGV 273
DB 246 EYKXGAVV 253

RESULT 33

Db 235 EBDQENKLYQKVEYQOEFTKVIATSKGANVPAM 271

RESULT 35
US-10-206-576-30
Sequence 30, Application US/10206576
Publication No. US20030017495A1
GENERAL INFORMATION:
APPLICANT: Choi et al.
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 497
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-R
COMPUTER: Dell Latitude
OPERATING SYSTEM: Windows 98
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/206,576
FILING DATE: 29-Jul-2002
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/071,035
FILING DATE: 1998-05-04
APPLICATION NUMBER: US 60/046,655
FILING DATE: 1997-05-16
APPLICATION NUMBER: US 60/044,031
FILING DATE: 1997-05-06
APPLICATION NUMBER: US 60/066,009
FILING DATE: 1997-11-14
ATTORNEY/AGENT INFORMATION:
NAME: Hyman, Mark J.
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB369P1D1
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 277 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 30:
US-10-206-576-30

Query Match 34.8%; Score 481; DB 12; Length 277;
Best Local Similarity 36.8%; Pred. No. 1.6e-35;
Matches 102; Conservative 62; Mismatches 101; Indels 12; Gaps 5;

QY 3 FGKINGICALASGIALAGCSNOSNEPAISKTPAQTIKVGWAGPEQVAEVAQVAK 62
Db 4 FSKIGIGIYLA-FTTAGCAGS-----VMDKTEVVKLGIVGTQNDWESVKRLKKK 56
QY YNLVELVEFNDYAMPNSAVSKGLDANMOKHPYLEKSOEKGNNLVIYGFVYPLA 122
Db 57 -NIDQIVERTDTPQNPAAAEKEIDLPFQDIPIDYNNKHG--TKLVSGNTNAPLG 114
QY 123 GYSTKITLVELDQATIAVNDPSNLARALLILEKQGIKDKDNTNLFSTLDIVENPK 182
Db 115 IYANKLKDITKIDGGEIAIPNDPTNGRALIILQTAGLIRKVDPAKQOLPTVSDITERK 174
QY 183 KLVKEVDTSVAARIDVDLAVVNNVAGVGLTASENGVE--DKDSPPYNIIVAR 239
Db 175 QLKTEIDATQTRALQDPDASVINGMAVDAGIYPPDKAIFLEFVNEKAPYNIIVAR 234
QY 240 ADNKDSFATDPEYKAYQTDVEAEAKQYDGVYIGW 276
Db 235 EBDQENKLYQKVEYQOEFTKVIATSKGANVPAM 271

RESULT 36
US-10-282-122A-54418
Sequence 54418, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Olsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forayth, R.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: Patentin version 3.1
SEQ ID NO 54418
LENGTH: 257
TYPE: PRT
ORGANISM: Campylobacter jejuni
US-10-282-122A-54418

Query Match 34.5%; Score 477.5; DB 12; Length 257;
Best Local Similarity 41.5%; Pred. No. 3e-35;
Matches 100; Conservative 50; Mismatches 78; Indels 13; Gaps 5;

QY 30 AISKTPAQTIKVGWAGPEQVAEVAQVAK--EKNLVELVEFNDYAMPNSAVSKGL 87
Db 15 SINAAALFTIIVAAATPVPH--AEILQVKNPDEKQYKLEIKFTDYVLPNLAVDNGEA 71
QY 88 DANMOKHPYLEKSOEKGNNLVIYGFVYPLAGSTIKTLNLEKQATIAVNDPS 147
Db 72 DANFQHTPYLEERNKNG--TKLKVAAIHIEPAVYSKYKSLDIDKGVKALPNDPT 130
QY 148 NLARALLILEKQGIKDKDNTNLFSTLDIVENPKQLVIEKVDTSVAARIDVDLAVN 207
Db 131 NESRALDIITAKGLVKEFDKA--LKTPLDIIDNPKKIKFVELKPAOLPRALNDVDPAVIN 188
QY 208 NNVAQVGLTASENGVEFEDCDSPYNIIVARNDNSKAIQDFVAKYQTDVEAEAKQ 267
Db 189 SNTALSNLNPADKDSVFIEDKESRYANILVVRVGHENDPKIKALQALQSDKI-----KQ 243
QY 268 F 268

Tue Jun 22 11:04:04 2004

us-10-018-672-2.rapb

Page 19

Db 244 F 244

```
RESULT 37
US-10-282-122A-52496
; Sequence 52496, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52496
; LENGTH: 270
; TYPE: PRT
; ORGANISM: Clostridium botulinum
US-10-282-122A-52496

Query Match 34.2%; Score 472.5; DB 12; Length 270;
Best Local Similarity 40.4%; Pred. No. 9.2e-35;
Matches 105; Conservative 54; Mismatches 86; Indels 13; Gaps 6;

QY 16 IALAGCSNOSNEPAISKTAQTIKYGVAGGEQVAEYAGQVAK--EKNLTVLELVFN 73
Db 19 VALTGGGS--SKEAKKKK---TIYVGATPEPH--AAILKKVZILKKKGYTLEIKFT 70
QY 74 DYAMPASVSKGELDANAMCHKPYLEKDSOEKGLNNLVIGNTFVYPLAGYSTKIKTINE 133
Db 71 DYTTPPTALQDGEIDANFYOHITPYLEEFYKKK--KTDLSTVYVHLEPBGVSVSKIKOLKE 129
QY 134 LKKGATITAVNDPSNLARALILILEKQGLIKLKDNTNLESTLIDYENKRYIKVETDTSV 193
Db 130 LKKGATITISPSDITNSRALIKLEKGIKKEGE--LWSKMDITKPKNKIKIELDAQ 187
QY 194 AARAIIDVDLAVNNYAGQVGLTASENGVEEDKDSFYVNIIVARADKDSKAIODFEK 253
Db 188 LPTLTDVDAVAANTNYAVANPLPKDALAIKSDSPYANIVVKTENKAAEYIKALDE 247
```

QY 254 AYOTDEVEAKKOPRODGI 273
Db 248 AINSEIKYIEQYKGAITL 267

```
RESULT 38
US-10-282-122A-52838
; Sequence 52838, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52838
; LENGTH: 270
; TYPE: PRT
; ORGANISM: Clostridium botulinum
US-10-282-122A-52838
```

```
Query Match 34.1%; Score 472; DB 12; Length 270;
Best Local Similarity 40.9%; Pred. No. 1e-34;
Matches 112; Conservative 46; Mismatches 108; Indels 8; Gaps 5;

QY 1 KMFCKINGI-CALASGIALAGCSNOSNEPAISKTAQTIKYGVAGGEQVAEYAGQVA 59
Db 1 MKKKRLTGLTLEVFPLTIGVCGSSKSKSEKTYVNDK---KTIYVGATPPVAGEILKVA--QPL 56
QY 60 KKKVNLTVLAVFENFYAUNDNSAVSKGELDANAMCHKPYLEKDSOEKGLNNLVIGNTFVY 119
Db 57 LKKKGYKLEIKFTYVTPNTLRNDREIDANFYOHITPYLEEFYKKK--KTELEAKRYKIYIS 115
QY 120 PLAGYSTKIKTINEKDGATTAVPNDPSNLARALILILEKQGLIKLKDNTNLESTLIDIVE 179
Db 116 PLATVSNKTKLEKDGATTAVPNDPTNERPALRLLEKSGILIKKGGDTL--TKGDIAE 173
QY 180 NPKVLYKVEDTSVAABAIDVDLAVNNYAGQVGLTASENGVEEDKDSFYVNIIVAR 239
```

Db 174 NKNKIKETIDAPQLPILNDVAVINTMYAEALTNLKSILIEDSDSYANVIANK 233
 QY 240 ADNKSQAIQDFKAYOTDEVEAEAKKQFQDGI 273
 Db 234 KEDKXAKIRALSELTSKEVKDFINEKXGVVI 267

RESULT 39

US-10-282-122A-54417
 ; Sequence 54417, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Treawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See file Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: Patent version 3.1
 ; SEQ ID NO 54417
 ; LENGTH: 256
 ; TYPE: PRT
 ; ORGANISM: Campylobacter jejuni
 ; US-10-282-122A-54417

Query Match 33.8%; Score 468; DB 12; Length 256;
 Best Local Similarity 42.2%; Pred. No. 2,2e-34;
 Matches 103; Conservative 52; Mismatches 83; Indels 6; Gaps 4;
 QY 31 ISKTAQTIKVGAGBQAVAEVAQVAKER-KYLTVELVEFNDYAMPNSAVSKGELDA 89
 Db 15 ISSFAQNTITIGATPFGSILEMDEPKNGYEL--KIVFSDYILPNPALEKEKEDA 72
 QY 90 NAMOHKRYLEKDSQEGUNLVYVGNTPFYPLAGYSTKIKTLNELKQDATTAVPDPBNL 149
 Db 73 NLIOHKEFSEYVWKKG-SYLIATTPVLIAPGVYSKIKLLENLISGGRVAIPYDATTNE 131
 QY 150 ARALILEKQGLIKLQNTNLFSTLIDIVENPKLVYKEVDTSVAAAIDVDLAVVNNN 209
 Db 132 SRALELEKAKLIEL--NKNTLTKPHIDINKPKKLFIEIKAAQCPALDVIDIATINSN 189

QY 210 YAGOVGLTASENGVEVEDEKSPYVNIIVARADNKSQAIQDFKAYOTDEVEAEAKKQF 269
 Db 190 PALGCLNLSKQTIREDKNSPYVYVVESEGNSEKTKYIDELIASDKFKALINHHX 249
 QY 270 DGYI 273
 Db 250 DILI 253

RESULT 40

US-09-815-242-11460
 ; Sequence 11460, Application US/09815242
 ; Patent No. US20020061569A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari L.
 ; APPLICANT: Zyskind, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Treawick, John D.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 ; TITLE OF INVENTION: Identification of Essential Genes in Prokaryotes
 ; FILE REFERENCE: ELITRA.011A
 ; CURRENT APPLICATION NUMBER: US/09/815,242
 ; CURRENT FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; NUMBER OF SEQ ID NOS: 14110
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 11460
 ; LENGTH: 271
 ; TYPE: PRT
 ; ORGANISM: Helicobacter pylori
 ; US-09-815-242-11460

Query Match 33.8%; Score 467.5; DB 9; Length 271;
 Best Local Similarity 42.1%; Pred. No. 2.7e-34;
 Matches 99; Conservative 46; Mismatches 85; Indels 5; Gaps 3;
 QY 39 IYGVWAGBQAVAEVAQVAKERYNTLVELVEFNDYAMPNSAVSKGELDANAYOKRYL 98
 Db 39 LKYGANPYVHAQIIQSDVDDLEKER-GIKLVIVSFYDVLNPLMLNDSLDANFYOHKPYL 97
 QY 99 EKDSQEGKGNLVYVGNTPFYPLAGYSTKIKTLNELKQDATTAVPDPBNL 158
 Db 98 DFNPDIRKM-RHVGANIHVEPLRFYSKTIIDINLKKGSIVAVPDPNPAQGGALLILIK 156
 QY 159 QGLIKLQNTNLFSTLIDIVENPKLVYKEVDTSVAAAIDVDLAVVNNNAGQVGLTA 218
 Db 157 QGLIKLQNTNLFSTLIDIVENPKLVYKEVDTSVAAAIDVDLAVVNNNAGQVGLTA 216
 QY 219 SENGVEVEDEKSPYVNIIVARADNKSQAIQDFKAYOTDEVEAEAKKQFQDGI 273
 Db 217 A--LFSQKDSFYANLNAEDNAQDEBALITLIEALQSEKTKRFILDTYKGAII 268

Search completed: June 16, 2004, 11:14:11
 Job time : 51 secs

Tue Jun 22 11:04:04 2004

us-10-018-672-2.rapb

Page 21



GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 16, 2004, 11:00:16 ; Search time 60 seconds
(without alignments)
1259.719 Million cell updates/sec

Title: US-10-018-672-2
Perfect score: 1383
Sequence: 1 MNFKINGICALASGIALAG.....TDEVEAKKQKGVKIGW 276

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_293an04:*

- 1: geneseqp19806:*
- 2: geneseqp19906:*
- 3: geneseqp20008:*
- 4: geneseqp20018:*
- 5: geneseqp20028:*
- 6: geneseqp20038:*
- 7: geneseqp20038:*
- 8: geneseqp20046:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1383	100.0	276	4	ABU20104 Moraxella
2	1368	98.9	276	6	AAO17582 Moraxella
3	1368	98.9	276	6	ABU35462 Protein e
4	1007.5	72.8	273	5	ABU76073 Haemophil
5	1006.5	72.8	273	4	AAU35474 Haemophil
6	1006.5	72.8	273	5	ABU76074 Haemophil
7	1006.5	72.8	273	5	ABU30316 Protein e
8	1004	72.6	272	5	ABU17813 H Influen
9	988	64.2	276	6	ABU39463 Protein e
10	793.5	57.4	271	6	ABU48271 Protein e
11	782.5	56.6	271	4	AAU34457 E. coli c
12	782.5	56.6	271	6	ABU15367 Protein e
13	780.5	56.4	271	6	ABU27902 Protein e
14	775.5	56.1	271	6	ABU31676 Protein e
15	767.5	55.5	272	6	ABU67576 Photornab
16	757.5	54.8	271	6	ABU50250 Protein e
17	757.5	54.8	271	6	ABU40467 Protein e
18	751.5	54.3	275	6	ABU49280 Protein e
19	743.5	53.8	240	4	AAU36063 Klebsiell
20	562	40.6	295	6	ABU19780 Protein e
21	551	39.8	262	6	ABU23447 Protein e
22	549.5	39.4	272	6	ABU22451 Protein e
23	545	39.4	256	6	ABU33293 Protein e
24	540	39.0	268	6	ABU21332 Protein e
25	527.5	38.1	270	6	ABU21415 Protein e

26	526.5	38.1	265	6	ABU23277 Protein e
27	509.5	36.8	286	5	ABU53625 Lactococ
28	506.5	36.6	261	6	ABU3768 Protein e
29	502.5	36.3	260	4	ABU36517 Pseudomon
30	502.5	36.3	260	6	ABU38920 Protein e
31	497.5	36.0	259	4	AAU33650 Pseudomon
32	497.5	36.0	259	6	ABU15545 Protein e
33	494	35.7	263	6	ABU25052 Protein e
34	488	35.3	255	6	ABU39703 Protein e
35	487	35.2	257	6	ABU1435 Protein e
36	481	34.8	277	2	AAU00024 Enterococ
37	481	34.8	277	5	ABP43243 E. faecali
38	481	34.8	277	6	ABU88271 E. faecali
39	481	34.8	277	6	ABU29285 Protein e
40	481	34.8	277	6	ABU13522 Enterococ
41	477.5	34.5	257	6	ABU26494 Protein e
42	472.5	34.2	270	6	ABU24572 Protein e
43	472	34.1	270	6	ABU24914 Protein e
44	469	33.8	256	6	ABU26493 Protein e
45	467.5	33.8	271	2	AAU30023 Expressed

ALIGNMENTS

RESULT 1
ID AAB20104 standard; protein; 276 AA.
XX
AC AAB20104;
XX
DT 23-APR-2001 (first entry)
XX
DE Moraxella catarrhalis BAS11 protein.
XX
KW BAS11; infection; otitis media; pneumonia; diagnosis; therapy;
KW antibacterial; antimicrobial; vaccine.
XX
OS Moraxella catarrhalis.
XX
PN MC2001.00637-A1.
XX
PD 04-JAN-2001.
XX
PF 23-JUN-2000; 2000MC-EP005552.
XX
PR 25-JUN-1999; 99GB-00014945.
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
P1 Thomas J;
XX
DR Weir; 2001-123013/13.
DR N-PSDB; AAF30040.
XX
PT New BAS11 polypeptides of Moraxella catarrhalis useful for diagnostic,
PT prophylactic and therapeutic purposes against microbial diseases,
PT preferably bacterial infections.
PS
PS Claim 1; Page 63; 79pp; English.
CC The present sequence is that of BAS11 protein from Moraxella
CC catarrhalis strain MC2931 (ATCC 43617), a causative agent of otitis media
CC in children and pneumonia in adults. The invention provides BAS11
CC polypeptides, and polynucleotides encoding them, as well as expression
CC vectors, host cells and methods for producing BAS11 polypeptides using
CC recombinant methods. Also claimed is a vaccine composition comprising a
CC BAS11 polypeptide, an immunogenic fragment of a BAS11 polypeptide, or
CC a polypeptide having at least 85% amino acid sequence identity to
CC BAS11, or comprising a polynucleotide encoding such a polypeptide. A
CC claimed method of diagnosing a Moraxella infection involves identifying a
CC BAS11 polypeptide or antibody. A claimed therapeutic composition useful
CC in treating humans with M. catarrhalis infection comprises at least 1

CC antibody directed against a BASH11 polypeptide. BASH11 polypeptides
CC also have utility in raising specific antibodies, and in screening for
CC antibacterial drugs
XX
SQ Sequence 276 AA;

Query Match 100.0%; Score 1383; DB 4; Length 276;
Best Local Similarity 100.0%; Pred. No. 1.4e-119;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNEFKINGICALAGSALAGCSNOSNEPAISKTAAQTIKVGWAGPEQAVAEVAGQVAK 60
DB 1 MNEFKINGICALAGSALAGCSNOSNEPAISKTAAQTIKVGWAGPEQAVAEVAGQVAK 60
QY 61 EKNLTVLVEFNDYMPNSAVSKGELDANAMQHKRYLEKDSQEGKLNLLVVGNTFYVP 120
DB 61 EKNLTVLVEFNDYMPNSAVSKGELDANAMQHKRYLEKDSQEGKLNLLVVGNTFYVP 120
QY 121 LAGSTKIKTLNELKDGATIAVPNDPSNLARALILEKQGLIKLKDNTNLFSTTLDIVEN 180
DB 121 LAGSTKIKTLNELKDGATIAVPNDPSNLARALILEKQGLIKLKDNTNLFSTTLDIVEN 180
QY 181 PKKLVIKEVDTSVAARAIDVDLAVVNNNYAGGVGLTASENGVFVEDKXSPYVNIIVARA 240
DB 181 PKKLVIKEVDTSVAARAIDVDLAVVNNNYAGGVGLTASENGVFVEDKXSPYVNIIVARA 240
QY 241 DNKSKAIDQFVAKYOTDEVEAEAKKQFKDGVYKGM 276
DB 241 DNKSKAIDQFVAKYOTDEVEAEAKKQFKDGVYKGM 276

RESULT 2
AA017582
ID AA017582 standard; protein; 276 AA.

XX AA017582;
XX 19-JUL-2002 (first entry)
XX
DE M catarrhalis MCA101896 protein SEQ ID NO: 44.

KW Moraxella; vaccine; respiratory tract infection; anti-inflammatory;
KM auditory; antibacterial; otitis media; sinusitis; pneumonia.

XX OS Moraxella catarrhalis.

XX PN WO200218595-A2.

XX PD 07-MAR-2002.

XX PF 28-AUG-2001; 2001WO-CA001221.

XX 28-AUG-2000; 2000US-0228294P.
PR 28-AUG-2000; 2000US-0228295P.
PR 28-AUG-2000; 2000US-0228296P.
PR 29-AUG-2000; 2000US-0228438P.
PR 29-AUG-2000; 2000US-0228439P.
PR 29-AUG-2000; 2000US-0228440P.
PR 29-AUG-2000; 2000US-0228441P.
PR 29-AUG-2000; 2000US-0228442P.
PR 29-AUG-2000; 2000US-0228443P.
PR 29-AUG-2000; 2000US-0228511P.
PR 29-AUG-2000; 2000US-0228512P.
PR 29-AUG-2000; 2000US-0228742P.
PR 29-AUG-2000; 2000US-0228743P.
PR 01-SEP-2000; 2000US-0229465P.
PR 01-SEP-2000; 2000US-0229474P.
PR 01-SEP-2000; 2000US-0229475P.
PR 01-SEP-2000; 2000US-0229478P.
PR 05-SEP-2000; 2000US-0229740P.
PR 05-SEP-2000; 2000US-0229803P.
PR 05-SEP-2000; 2000US-0229804P.
PR 05-SEP-2000; 2000US-0229805P.

N-2

PR 05-SEP-2000; 2000US-0229806P.
PR 05-SEP-2000; 2000US-0229809P.
PR 05-SEP-2000; 2000US-0229811P.
PR 06-SEP-2000; 2000US-0230214P.
PR 06-SEP-2000; 2000US-0230250P.
PR 06-SEP-2000; 2000US-0230252P.
XX
XX (AVET) AVENTIS PASTEUR LTD.
XX
XX Loosmore S, Wang J, Bradley B, Ochs M, Yang Y;
XX WPI; 2002-404555/43.
XX DR N-PSDB; AAL46514.

PT Moraxella polypeptide and polynucleotides useful as vaccine for
PT immunizing a host e.g. humans against disease e.g. otitis media,
PT pneumonia, caused by infection of the bacteria.

PS Claim 28; Fig 43; 277pp; English.

CC The present invention provides the protein and coding sequences of
CC proteins from Moraxella catarrhalis. These can be used to produce
CC vaccines which protect against M. catarrhalis infection, which can cause
CC otitis media, respiratory infection, sinusitis, and pneumonia. The
CC present sequence is a protein of the invention

XX SQ Sequence 276 AA;

Query Match 98.3%; Score 1368; DB 5; Length 276;
Best Local Similarity 99.3%; Pred. No. 3.5e-118;
Matches 274; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MNEFKINGICALAGSALAGCSNOSNEPAISKTAAQTIKVGWAGPEQAVAEVAGQVAK 60
DB 1 MNEFKINGICALAGSALAGCSNOSNEPAISKTAAQTIKVGWAGPEQAVAEVAGQVAK 60
QY 61 EKNLTVLVEFNDYMPNSAVSKGELDANAMQHKRYLEKDSQEGKLNLLVVGNTFYVP 120
DB 61 EKNLTVLVEFNDYMPNSAVSKGELDANAMQHKRYLEKDSQEGKLNLLVVGNTFYVP 120
QY 121 LAGSTKIKTLNELKDGATIAVPNDPSNLARALILEKQGLIKLKDNTNLFSTTLDIVEN 180
DB 121 LAGSTKIKTLNELKDGATIAVPNDPSNLARALILEKQGLIKLKDNTNLFSTTLDIVEN 180
QY 181 PKKLVIKEVDTSVAARAIDVDLAVVNNNYAGGVGLTASENGVFVEDKXSPYVNIIVARA 240
DB 181 PKKLVIKEVDTSVAARAIDVDLAVVNNNYAGGVGLTASENGVFVEDKXSPYVNIIVARA 240
QY 241 DNKSKAIDQFVAKYOTDEVEAEAKKQFKDGVYKGM 276
DB 241 DNKSKAIDQFVAKYOTDEVEAEAKKQFKDGVYKGM 276

RESULT 3
ABU35462
ID ABU35462 standard; protein; 276 AA.

XX ABU35462;

XX 19-JUN-2003 (first entry)

XX Protein encoded by Prokaryotic essential gene #20989.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX OS Moraxella catarrhalis.

XX PN WO200277183-A2.

XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002WO-US009107.

21-MAR-2001; 2001US-00815242.
06-SEP-2001; 2001US-00948993.
25-OCT-2001; 2001US-0342923P.
08-FEB-2002; 2002US-00072851.
06-MAR-2002; 2002US-0362699P.
(ELIT-) ELITRA PHARM INC.
Wang L, Zamudio C, Malone C, Haselbeck R, Ohleen KL, Zyskind JW,
Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH,
WPI; 2003-029926/02.
N-PSDB; ACA39332.
New antisense nucleic acids, useful for identifying proteins or screening
for homologous nucleic acids required for cellular proliferation to
isolate candidate molecules for rational drug discovery programs.
Claim 25; SEQ ID NO 63386; 1766bp; English.
The invention relates to an isolated nucleic acid comprising any one of
the 623 antisense sequences given in the specification where expression
of the nucleic acid inhibits proliferation of a cell. Also included are:
(1) a vector comprising a promoter operably linked to the nucleic acid
encoding a polypeptide whose expression is inhibited by the antisense
nucleic acid; (2) a host cell containing the vector; (3) an isolated
polypeptide or its fragment whose expression is inhibited by the
antisense nucleic acid; (4) an antibody capable of specifically binding
the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
proliferation; (7) identifying a compound that influences the activity of
the gene product or that has an activity against a biological pathway
required for proliferation, or that inhibits cellular proliferation; (8)
identifying a gene required for cellular proliferation or the biological
pathway in which a proliferation-required gene or its gene product lies
or a gene on which the test compound that inhibits proliferation of an
organism acts; (9) manufacturing an antidiabetic; (10) profiling a
compound's activity; (11) a culture comprising strains in which the gene
product is overexpressed or underexpressed; (12) determining the extent
to which each of the strains is present in a culture or collection of
strains; or (13) identifying the target of a compound that inhibits the
proliferation of an organism. The antisense nucleic acids are useful for
identifying proteins or screening for homologous nucleic acids required
for cellular proliferation to isolate candidate molecules for rational
drug discovery programs, or for screening homologous nucleic acids
required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
K. pneumoniae or *P. aeruginosa*. The present sequence is encoded by one of
the target prokaryotic essential genes. Note: The sequence data for this
patent did not form part of the printed specification, but was obtained
in electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences
Sequence 276 AA;
Query Match 98.9%; Score 1368; DB 6; Length 276;
Best Local Similarity 99.3%; Pred. No. 3.5e-118;
Matches 274; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

241 DNKSKAIDPFVAKYOTDEVEAEAKKOFKGVTKGM 276
241 DNKSKAIDPFVAKYOTDEVEAEAKKOFKGVTKGM 276
RESULT 4
ABB76073
ID ABB76073 standard; protein; 273 AA.
AC ABB76073;
DT 15-JUL-2002 (first entry)
XX Haemophilus influenzae BASB202 protein.
DE Haemophilus influenzae BASB202 protein.
XX BASB202; NTH; infection, vaccine; genetic immunisation; auditory;
KM antiinflammatory; antibacterial; immunostimulant; otitis media.
XX Haemophilus influenzae.
XX Location/Qualifiers
FH 1..15
FT /note= "potential T-helper cell epitope"
FT 21..28
FT /note= "potential B-cell epitope"
FT 34..42
FT /note= "potential T-helper cell epitope"
FT 43..47
FT /note= "potential B-cell epitope"
FT 56..60
FT /note= "potential B-cell epitope"
FT 66..74
FT /note= "potential T-helper cell epitope"
FT 69..74
FT /note= "potential B-cell epitope"
FT 78..84
FT /note= "potential B-cell epitope"
FT 89..97
FT /note= "potential T-helper cell epitope"
FT 93..104
FT /note= "potential B-cell epitope"
FT 107..126
FT /note= "potential T-helper cell epitope"
FT 121..127
FT /note= "potential B-cell epitope"
FT 129..134
FT /note= "potential B-cell epitope"
FT 141..148
FT /note= "potential B-cell epitope"
FT 150..163
FT /note= "potential T-helper cell epitope"
FT 176..182
FT /note= "potential B-cell epitope"
FT 181..189
FT /note= "potential T-helper cell epitope"
FT 200..208
FT /note= "potential T-helper cell epitope"
FT 224..228
FT /note= "potential B-cell epitope"
FT 229..240
FT /note= "potential T-helper cell epitope"
FT 236..244
FT /note= "potential B-cell epitope"
FT 249..256
FT /note= "potential B-cell epitope"
FT 259..267
FT /note= "potential B-cell epitope"
WO200224729-A2.
28-MAR-2002.

PF 18-SEP-2001; 2001WO-EP010979.
 XX
 XX 19-SEP-2000; 2000GB-00022992.
 XX
 PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 PI Thomard J;
 XX
 XX WPI, 2002-383180/41.
 DR N-PSDB; ABL56953, ABL56954, ABL56956.
 XX
 PT New isolated BASB202 polypeptide of non-typeable Haemophilus influenzae
 PT useful for diagnosing a disease and in generating an immune response in
 PT an animal.
 PS
 PS Claim 3; Page 75; 90pp; English.
 XX
 XX The present sequence is the protein sequence for the BASB202 protein of
 CC non-typeable Haemophilus influenzae (NTHi) strain 3224A (ATCC PT-1816)
 CC isolated from an otitis media patient in the USA. NTHi strain 3219C
 CC isolated from an otitis media patient in the USA. NTHi strain 810956
 CC isolated from a meningitis patient in the Netherlands, and NTHi strain
 CC A40184, a carrier strain from the Netherlands. Variability analysis of
 CC BASB202 protein among NTHi strains revealed 99-100% sequence identity.
 CC The invention provides BASB202 polypeptides and polynucleotides, vectors,
 CC host cells, and methods for producing the polypeptides by recombinant
 CC methods. Claimed vaccine compositions comprise a BASB202 polypeptide or
 CC polynucleotide. A claimed method of diagnosing NTHi infection involves
 CC identifying a BASB202 polypeptide or an antibody that is immunospecific
 CC for the polypeptide. A claimed therapeutic composition useful for
 CC treating humans with NTHi disease comprises an antibody directed against
 CC a BASB202 polypeptide. NTHi diseases include otitis media, pneumonia,
 CC sinusitis, nosocomial infections, invasive disease, chronic otitis media
 CC with hearing loss, fluid accumulation in the middle ear, auditory nerve
 CC damage, delayed speech learning, infection of the upper respiratory
 CC tract, and inflammation of the middle ear. B-cell epitopes and T-helper
 CC cell epitopes from BASB202 may also be useful in vaccine compositions
 XX
 XX Sequence 273 AA;
 SQ
 Query Match 72.8%; Score 1007.5; DB 5; Length 273;
 Best Local Similarity 71.7%; Pred. No. 8.7e-85;
 Matches 198; Conservative 27; Mismatches 48; Indels 3; Gaps 1;

DE Haemophilus influenzae cellular proliferation protein #115.
 XX
 XX Antisense; prokaryotic cellular proliferation protein; antibiotic;
 KW antibacterial; drug design.
 XX
 XX Haemophilus influenzae.
 OS
 XX WO200170955-A2.
 XX
 XX 27-SEP-2001.
 XX
 XX 21-MAR-2001; 2001WO-US009180.
 XX
 XX 21-MAR-2000; 2000US-0191078P.
 XX 23-MAY-2000; 2000US-0206848P.
 XX 26-MAY-2000; 2000US-0207727P.
 XX 23-OCT-2000; 2000US-0242578P.
 XX 27-NOV-2000; 2000US-0253625P.
 XX 22-DEC-2000; 2000US-0257931P.
 XX 16-FEB-2001; 2001US-0269308P.
 XX
 XX (ELITRA) PHARM INC.
 XX
 XX Haseelbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX
 XX WPI; 2001-611495/70.
 DR N-PSDB; AAS53333.
 XX
 XX New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids.
 PT
 XX Example 3; SEQ ID NO 11067; 511pp; English.
 PS
 XX The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the genes,
 CC their use in the discovery of novel antibiotics, the essential genes
 CC themselves and the encoded proteins. The prokaryotes used are Escherichia
 CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
 CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
 CC useful for the identification of potential new targets for antibiotic
 CC development. The antisense nucleic acids can also be used to identify
 CC proteins used in proliferation, to express these proteins, and to obtain
 CC antibodies capable of binding to the expressed proteins. The proteins can
 CC be used to screen compounds in rational drug discovery programmes. The
 CC antisense nucleic acid sequence is also useful to screen for homologous
 CC nucleic acids which are required for cell proliferation in a wide variety
 CC of organisms. The present sequence represents an essential prokaryotic
 CC cellular proliferation protein. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 CC
 XX Sequence 273 AA;
 SQ
 Query Match 72.8%; Score 1006.5; DB 4; Length 273;
 Best Local Similarity 71.4%; Pred. No. 1.1e-84;
 Matches 197; Conservative 28; Mismatches 48; Indels 3; Gaps 1;

Db 178 PKKLITEVDTSVAARALDDVDLAVNNNTYAQGVGLNAQDDGVFEVDKDSPPYNIIVSRT 237
QY 241 DNKDSKAIDPFVAYQTDVEVEAEAKKQFKDGVKGM 276
Db 238 DNKDSKAQVDFIKSYOTEVEVYQEAQKHFKDGVVKGW 273

RESULT 6
ABU30316
ID ABB76074 standard; protein; 273 AA.

AC ABB76074;

XX 15-JUN-2002 (first entry)

XX Haemophilus influenzae BASB202 protein.

XX BASB202; NTHi; infection; vaccine; genetic immunisation; auditory;
XX antiinflammatory; antibacterial; immunostimulant; otitis media.

XX Haemophilus influenzae.

XX Key Location/Qualifiers

XX Region 1..15 /note="potential T-helper cell epitope"

XX Region 21..28 /note="potential B-cell epitope"

XX Region 34..42 /note="potential T-helper cell epitope"

XX Region 43..47 /note="potential B-cell epitope"

XX Region 56..60 /note="potential B-cell epitope"

XX Region 66..74 /note="potential B-cell epitope"

XX Region 69..74 /note="potential T-helper cell epitope"

XX Region 78..84 /note="potential B-cell epitope"

XX Region 89..97 /note="potential B-cell epitope"

XX Region /note="potential T-helper cell epitope"

XX Region 93..104 /note="potential B-cell epitope"

XX Region 107..126 /note="potential T-helper cell epitope"

XX Region 121..127 /note="potential B-cell epitope"

XX Region 129..134 /note="potential B-cell epitope"

XX Region 141..148 /note="potential B-cell epitope"

XX Region 150..169 /note="potential B-cell epitope"

XX Region 176..182 /note="potential T-helper cell epitope"

XX Region 181..189 /note="potential B-cell epitope"

XX Region 200..208 /note="potential T-helper cell epitope"

XX Region 224..228 /note="potential T-helper cell epitope"

XX Region 229..240 /note="potential B-cell epitope"

XX Region 236..244 /note="potential T-helper cell epitope"

XX Region 249..256 /note="potential B-cell epitope"

XX Region 259..267 /note="potential B-cell epitope"

XX 18-SEP-2001; 2001WO-EP010979.
XX 19-SEP-2000; 2000GB-00022992.
XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX Thomas J;
XX WPI; 2002-383180/41.
XX N-PSDB; ABL56955.
XX New isolated BASB202 polypeptide of nontypeable Haemophilus influenzae
XX useful for diagnosing a disease and in generating an immune response in
XX an animal.
XX Claim 3; Page 76; 90pp; English.

XX The present sequence is the protein sequence for the BASB202 protein of
XX nontypeable Haemophilus influenzae (NTHi) strain 27N116791N1 isolated
XX from a cystic fibrosis patient in Denmark. Variability analysis of
XX BASB202 protein among different NTHi strains revealed 99-100% sequence
XX identity. The invention provides BASB202 polypeptides and
XX polynucleotides, vectors, host cells, and methods for producing the
XX polypeptides by recombinant methods. Claimed vaccine compositions
XX comprise a BASB202 polypeptide or polynucleotide. A claimed method of
XX diagnosing NTHi infection involves identifying a BASB202 polypeptide or
XX an antibody that is immunospecific for the polypeptide. A claimed
XX therapeutic composition useful for treating humans with NTHi disease
XX comprises an antibody directed against a BASB202 polypeptide. NTHi
XX diseases include otitis media, pneumonia, sinusitis, nosocomial
XX infections, invasive disease, chronic otitis media with hearing loss,
XX fluid accumulation in the middle ear, auditory nerve damage, delayed
XX speech learning, infection of the upper respiratory tract, and
XX inflammation of the middle ear. B-cell epitopes and T-helper cell
XX epitopes from BASB202 may also be useful in vaccine compositions

XX Sequence 273 AA:

Query Match 72.8%; Score 1006.5; DB 5; Length 273;
Best Local Similarity 71.7%; Pred. No. 1.1e-84;
Matches 198; Conservative 26; Mismatches 49; Indels 3; Gaps 1;

QY 1 MNFGKINGICALASGIALACCSNOSNEPAISTTAQTIKVGWAGPEQAAVEVAGQVAK 60
Db 1 MKLKQPLAIRAASALVLTGCKEDKPEAA---AAPKIKVGVMSGEHVAIEAAVAK 57
QY 61 EKYNLTVELPEFNDYAMPNSAVSKGELDANAMOHKPYLEKQSEKGLANTVYGNTPVP 120
Db 58 EKYGLDQVFEFNDYALPNEAVSKGDLNANAMOHKPYLEKQSEKGLANTVYGNTPVP 117
QY 121 LAGYSTKIKTLNELKQDATAVPNDPSNLARALILEKQGLIKLKQNTNLFSTLDIVEN 180
Db 118 LAGYSTKIKTLNELKQDATAVPNDPSNLARALILEKQGLIKLKQNTNLFSTLDIVEN 177
QY 181 PKKLIVKEVDTSVAARALDDVDLAVNNNTYAQGVGLTASNGVFEVDKDSPPYNIIVARA 240
Db 178 PKKLITEVDTSVAARALDDVDLAVNNNTYAQGVGLTASNGVFEVDKDSPPYNIIVSRT 237
QY 241 DNKDSKAIDPFVAYQTDVEVEAEAKKQFKDGVKGM 276
Db 238 DNKDSKAQVDFIKSYOTEVEVYQEAQKHFKDGVVKGW 273

RESULT 7
ABU30316
ID ABU30316 standard; protein; 273 AA.

XX ABU30316;

XX 19-JUN-2003 (first entry)

XX Protein encoded by Prokaryotic essential gene #15843

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX Haemophilus influenzae.
XX WO200277183-A2.
XX PD 03-OCT-2002.
XX 21-MAR-2002; 2002MO-US009107.
XX 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Hasebeck R, Ohlsen KJ, Zykkind JW,
XX Wall D, Trawick J, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI: 2003-029926/02.
XX N-PSDB; ACA34186.
XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX Claim 25; SEQ ID NO 58240; 1766bp; English.
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed; (12) determining the extent
XX to which each of the strains is present in a culture or collection of
XX strains; or (13) identifying the target of a compound that inhibits the
XX proliferation of an organism. The antisense nucleic acids are useful for
XX identifying proteins or screening for homologous nucleic acids required
XX for cellular proliferation to isolate candidate molecules for rational
XX drug discovery programs, or for screening homologous nucleic acids
XX required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
XX *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
XX the target prokaryotic essential genes. Note: the sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pat_sequences
XX
XX Sequence 273 AA:
XX
XX Query Match 72.8%; Score 1006.5; DB 6; Length 273;
XX Best Local Similarity 71.4%; Pred. No. 1.1e-84;
XX Matches 197; Conservative 28; Mismatches 48; Indels 3; Gaps 1;
OY 1 MNFGKINGICATLASGIALACSNOSNEPAISKTAOTIKVGVWAGPEQAVAEVAGQVAK 60
DB 1 MKLQQLPAIRTAISALVLTGCKEDKPEAA---AALPKIKVGVWAGPEQAVAEVAGQVAK 57

OY 61 EKVNLVEVEFNDYAMPNSAVSGEIDANAMOKPKYLEKSOEKGILNVLVIGNTFVYP 120
DB 58 EKVGLDVQFVEFNDYALPNBAVSKGDLNANAMQHPFLDEDAKQNLNVLVIGNTFVYP 117
OY 121 LAGYSTIKITLNEKDGATTAVPNDPSNLRALILIEKQGLITLXNTNLFTSTLTVEN 180
DB 118 LAGYSKIKXVNELODAKVVPPDPTNRGAILLEKQGLITLXNTNLFTSTLTVEN 177
OY 181 PKLIVKEVDTSYAARAIDVDLAVVNNYAGVGITLASENGFVEDEKSPYNIIVARA 240
DB 178 PKLITEVDTSYAARAIDVDLAVVNNYAGVGITLASENGFVEDEKSPYNIIVARA 237
OY 241 DNKSKAIDPFVAKYQTEVEAEAKKQKGVITKM 276
DB 238 DNKSKAVQDFIKSYQTEVEAEAKKQKGVITKM 273
RESULT 8
AA017813
ID AA017813 standard; protein: 272 AA.
AC AA017813;
XX 05-AUG-2002 (first entry)
DT H influenzae BVH-NTH112 protein SEQ ID NO: 24.
XX
XX Haemophilus influenzae infection; BVH-NTH11; otitis media; BVH-NTH12;
XX sinusitis; bronchitis; pneumonia; meningitis; bacteraemia; BVH-NTH13;
XX BVH-NTH14; BVH-NTH15; BVH-NTH16; BVH-NTH17; BVH-NTH18; BVH-NTH19;
XX BVH-NTH10; BVH-NTH11; BVH-NTH12; anti-inflammatory; auditory;
XX antibiotic; vaccine.
XX
XX Haemophilus influenzae.
XX WO200228889-A2.
XX PD 11-APR-2002.
XX 02-OCT-2001; 2001WO-CA001402.
XX 02-OCT-2000; 2000US-0236712P.
XX (SHIR-) SHIRE BIOCHEM INC.
XX Hamel J, Couture F, Brodeur BR, Martin D, Ouellet C, Tremblay M;
XX Charbonneau A, Vaysier C;
XX WPI: 2002-435325/46.
XX N-PSDB; AAL46811.
XX Novel isolated Haemophilus influenzae polypeptides BVH-NTH11-12, useful
XX for inducing protective immune responses against H. influenzae in animals
XX and for treating otitis media, sinusitis, bronchitis and pneumonia.
XX Claim 17; Fig 24; 58pp; English.
XX The present invention provides the protein and coding sequences of
XX Haemophilus influenzae BVH-NTH11-12. The sequences can be used in the
XX production of a vaccine to protect against, and in the diagnosis of, H.
XX influenzae infection, which can lead to otitis media, sinusitis,
XX bronchitis, pneumonia, meningitis and bacteraemia. The present sequence
XX is a protein of the invention
XX
XX Sequence 272 AA:
XX
XX Query Match 72.6%; Score 1004; DB 5; Length 272;
XX Best Local Similarity 71.4%; Pred. No. 1.8e-84;
XX Matches 197; Conservative 28; Mismatches 47; Indels 4; Gaps 1;
OY 1 MNFGKINGICATLASGIALACSNOSNEPAISKTAOTIKVGVWAGPEQAVAEVAGQVAK 60
DB 1 MKLQQLPAIRTAISALVLTGCKEDKPEAA---AALPKIKVGVWAGPEQAVAEVAGQVAK 56

QY 61 EKXNLTVELVEFNDYAMPNSAVSKGELDANAMQHKPYLEKDSQEKLNNTLVGTFTYYP 120
 DB 57 EKKGLVQVEFENDVALPNEAVAKGDLDMAMQHKPYLEDEDAKAKNLNLTIVGNTFYYP 116
 QY 121 LAGYSTRKIKTLNELKDGATIAVPNDPSNLRALILILEKQGLIKLXONTNLFSTTLDIYEN 180
 DB 117 LAGYSKKIKNNVELDQGAIVVPNDPTNGRALLILEKQGLIKLMDANNLTSTVDIYEN 176
 QY 181 PKKLVYKEVDTSVAAARALDDVLAANNVNNYAGQVGLTASNGVFEVDKSPYVNTIIVARA 240
 DB 177 PKKLTITEVDTSVAAARALDDVLAANNVNNYAGQVGLTASNGVFEVDKSPYVNTIIVARA 236
 QY 241 DNKDSKAIODFVKAYOTDEVEAEAKKQFQKGVYKGM 276
 DB 237 DNKDSKAVDDFVKASYOTEVEVYQEAQKHFQGVVYKGM 272

RESULT 9
 ABU39463
 ID ABU39463 standard; protein; 276 AA.
 XX
 AC ABU39463;

XX 19-JUN-2003 (first entry)
 DT Protein encoded by Prokaryotic essential gene #24990.
 XX
 DE Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Pasteurella multocida.
 OS

XX WO200277183-A2.
 PN

XX 03-OCT-2002.
 PD

XX 21-MAR-2002; 2002MO-US009107.
 PF

XX 21-MAR-2001; 2001US-00815242.
 PR

XX 06-SEP-2001; 2001US-00948993.
 PR

XX 25-OCT-2001; 2001US-0342923P.
 PR

XX 08-FEB-2002; 2002US-00072851.
 PR

XX 06-MAR-2002; 2002US-0362699P.
 PR

XX (ELIT-) ELITRA PHARM INC.
 PA

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI. 2003-029926/02.
 DR N-PSDB; ACA43333.

XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PI isolate candidate molecules for rational drug discovery programs.
 XX

XX Claim 25; SEQ ID NO 67387; 1766bp; English.
 PS

XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies

CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC parent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 CC

XX Sequence 276 AA;

XX Query Match 64.2%; Score 888; DB 6; Length 276;
 XX Best local similarity 60.2%; Pred. No. 1e-73;
 XX Matches 168; Conservative 55; Mismatches 50; Indels 6; Gaps 3;

QY 1 MNEFKINGICALAGIALAGCSNQSNEPAIS--KTAAGTIRYGVWAGPECAVAEYAGQ 57
 DB 1 MKLTKLFGLATIVSAVALAGC--KDKPFAAARPOEPANAKLTGVWGTGAQVTEVAAK 58

QY 58 VAKEXNLTVELVEFNDYAMPNSAVSKGELDANAMQHKPYLEKDSQEKLNNTLVGTFTYYP 117
 DB 59 IAKEXNIDVAVLVEFEFTYTPNDALTKGDLDNAFQHKPYMDXEVORGY-KLAIVGNTF 117

QY 118 VYPLAGYSTKIKTLNELKDGATIAVPNDPSNLRALILILEKQGLIKLXONTNLFSTTLDI 177
 DB 118 VFPIAASKKIKNNVELDQGAIVVPNDPTNGRALLILEKQGLIKLMDANNLTSTVDIYEN 177

QY 178 VENPKLVYKEVDTSVAAARALDDVLAANNVNNYAGQVGLTASNGVFEVDKSPYVNTIIV 237
 DB 178 IENPKNLQIKEVEGSLIPRLMDVDCFALINNNAVVOGGLTAEKDFVEVDKSPYVNTIIV 237

QY 238 ABADNKSRAIODFVKAYOTDEVEAEAKKQFQKGVYKGM 276
 DB 238 SREDNKNBAIKDFVAFOTEVEVYQEAQKHFQGVVYKGM 276

RESULT 10
 ABU48271

ID ABU48271 standard; protein; 271 AA.
 XX

XX ABU48271;
 AC

XX 19-JUN-2003 (first entry)
 DT

XX Protein encoded by Prokaryotic essential gene #3799.
 DE

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX

XX *Salmonella typhi*.
 OS

XX WO200277183-A2.
 PN

XX 03-OCT-2002.
 PD

XX 21-MAR-2002; 2002MO-US009107.
 PF

XX 21-MAR-2001; 2001US-00815242.
 PR

XX 06-SEP-2001; 2001US-00948993.
 PR

XX 25-OCT-2001; 2001US-0342923P.
 PR

XX 08-FEB-2002; 2002US-00072851.
 PR

XX 06-MAR-2002; 2002US-0362699P.
 PR

XX (ELIT-) ELITRA PHARM INC.
 PA Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI

PI Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
DR N-PSDB; ACAS2141.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
PS Claim 25; SEQ ID NO 76195; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 271 AA;
SQ
Query Match 57.4%; Score 793.5; DB 6; Length 271;
Best Local Similarity 56.5%; Pred. No. 5,8e-65;
Matches 156; Conservative 44; Mismatches 65; Indels 11; Gaps 3;
3 FGIKINGICAIASGIALAGCSNOSNEPAISKTAQTITKGVAMGPBOVAEVAEYGVAXEK 62
5 FKTFAAVGALIGSLALAGCGODEKDP-----NHIKGVAVGABQOVAEVAOKVAKEX 56
DB YNLTVLVEFNDYAMPNSAVSKGELDANAMQHPYLEKSOEKGNNLVIYVNTFYVPIA 122
63 YNLTVLVEFNDYAMPNSAVSKGELDANAMQHPYLEKSOEKGNNLVIYVNTFYVPIA 122
DB YGLDVELVTFNDYVLPREALSKGIDIDANAFQHKPYIDQOKRGY-KLVAVGNTFYVPIA 115
57 YGLDVELVTFNDYVLPREALSKGIDIDANAFQHKPYIDQOKRGY-KLVAVGNTFYVPIA 115
QY 123 GYSTKIKTINELKDGATIAVPNDPSNLARALILBKQGLIKKDKNTNLTSTLDIYENPK 182
DB 116 GYSKRTIKSIDELKDSQVAVPNDPTNLGRSLLLQVGLIKKDKDVSLLPTSLDIYENPK 175
QY 183 KLVYKRVDSVBARAIDD--VDLAIVNNNYAGOVGTASBENFVEDDQSPVNIIVARA 240
DB 176 NLKIVLELAPQLEFRSDDAQIALAVNTTYSQIGLTPKDGILFVBDKDSFVNLIVITRE 235
QY 241 DNEKSRADIDFVAYQYOTDEVEAEAKKQFKDGVYIKGW 276
DB 236 DNEKDAENVKKFVAYQSDVEYEAANKVFNGAVKGW 271
RESULT 11
AAU34457
ID AAU34457 standard; protein; 271 AA.

XX AAU34457;
XX AC
XX 14-FEB-2002 (first entry)
XX DT
XX DE *E. coli* cellular proliferation protein #38.
XX KM
XX Antisense; prokaryotic cellular proliferation protein; antibiotic;
XX antibiotic; drug design.
XX OS
XX *Escherichia coli*.
XX PN
XX MO200170955-A2.
XX PD
XX 27-SEP-2001.
XX PF
XX 21-MAR-2001; 2001MO-US009180.
XX PR
XX 21-MAR-2000; 2000US-0191078P.
XX PR 23-MAY-2000; 2000US-0206848P.
XX PR 26-MAY-2000; 2000US-0207727P.
XX PR 23-OCT-2000; 2000US-0242578P.
XX PR 27-NOV-2000; 2000US-0236252P.
XX PR 22-DEC-2000; 2000US-0257931P.
XX PR 16-FEB-2001; 2001US-0269308P.
XX PA
XX (ELIT-) ELITRA PHARM INC.
XX PI
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GU;
XX Yamamoto RT, Xu HH;
XX WPI; 2001-611495/70.
XX DR N-PSDB; AAS52316.
XX PT
XX New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids.
XX
XX Example 3; SEQ ID NO 10050; 51pp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
XX prokaryotic cellular proliferation, their use in identifying the genes,
XX their use in the discovery of novel antibiotics, the essential genes
XX themselves and the encoded proteins. The prokaryotes used are *Escherichia*
XX *coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,
XX *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also
XX useful for the identification of potential new targets for antibiotic
XX development. The antisense nucleic acids can also be used to identify
XX proteins used in proliferation, to express these proteins, and to obtain
XX antibodies capable of binding to the expressed proteins. The proteins can
XX be used to screen compounds in rational drug discovery programmes. The
XX antisense nucleic acid sequence is also useful to screen for homologous
XX nucleic acids which are required for cell proliferation in a wide variety
XX of organisms. The present sequence represents an essential prokaryotic
XX cellular proliferation protein. Note: The sequence data for this patent
XX did not form part of the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 271 AA;
SQ
Query Match 56.4%; Score 782.5; DB 4; Length 271;
Best Local Similarity 55.1%; Pred. No. 6.1e-64;
Matches 152; Conservative 46; Mismatches 67; Indels 11; Gaps 3;
3 FGIKINGICAIASGIALAGCSNOSNEPAISKTAQTITKGVAMGPBOVAEVAEYGVAXEK 62
5 FKTFAAVGALIGSLALAGCGODEKDP-----NHIKGVAVGABQOVAEVAOKVAKEX 56
DB YNLTVLVEFNDYAMPNSAVSKGELDANAMQHPYLEKSOEKGNNLVIYVNTFYVPIA 122
QY 63 YNLTVLVEFNDYAMPNSAVSKGELDANAMQHPYLEKSOEKGNNLVIYVNTFYVPIA 122
DB 57 YGLDVELVTFNDYVLPREALSKGIDIDANAFQHKPYIDQOKRGY-KLVAVGNTFYVPIA 115
QY 123 GYSTKIKTINELKDGATIAVPNDPSNLARALILBKQGLIKKDKNTNLTSTLDIYENPK 182


```

DB 116 GSKTKIKSLDELQDSQVAVPNDPTNLGRSLTLQVAGIKLKDQVGLPTVLADVENPK 175
QY 183 KLVIVKEDVTSVAARAIDD--VDLAVNNNNYAGQVGTASENGFVEDKDSPPVNIIVARA 240
DB 176 NLKIVLEAPQLPRSLDDDAQIALAVINTTYSQIGLTPKDGIFVEDKESPPVNIIVTRE 235
QY 241 DNKDSKAIQDFVKAYQTDVEVAEAKKQFQGVYKGM 276
DB 236 DNKDAENVKKFFVQAYQSDVEVEAANKVFNGAVKGM 271

RESULT 12
ABU15367
ID ABU15367 standard; protein; 271 AA.
XX
AC ABU15367;
XX
DE 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #694.
XX
OS Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Escherichia coli.
XX
FN WO20027183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WC-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR MPI: 2003-029926/02.
DR N-PSDB; ACA19237.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 43291; 1766bp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the

```

```

CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 271 AA;
XX
Query Match 56.6%; Score 782.5; DB 6; Length 271;
Best Local Similarity 55.1%; Pred. No. 6, 1e-64;
Matches 152; Conservative 46; Mismatches 67; Indels 11; Gaps 3;
QY 3 FGIKNGICATASGIALAGCSQSNVEPAISKTAQTIKGVYMGPEQVAEVAQVAKEX 62
DB 5 FKTPAAVGAIGSLALVGCQDEKDP-----NHIKGVIVAEQVAVAKAKDX 56
QY 63 YNLTVLEVEFNDYAMPNSAVSKGELDANAMQHKPYLEKDSQEKGLNVLIVGNTFVYPLA 122
DB 57 YGLDVELVTFNDYVLEFMBALSKGIDIDANAFQHKPYLDQQLKDRGY-KLVAVGNTFVYPIA 115
QY 123 GYSTKIKITNLEKQATIAVPNDPSNLARALLILEKQGLIKLQNTNLFSTLLDIYENPK 182
DB 116 GYSKTKIKSLDELQDSQVAVPNDPTNLGRSLTLQVAGIKLKDQVGLPTVLADVENPK 175
QY 183 KLVIVKEDVTSVAARAIDD--VDLAVNNNNYAGQVGTASENGFVEDKDSPPVNIIVARA 240
DB 176 NLKIVLEAPQLPRSLDDDAQIALAVINTTYSQIGLTPKDGIFVEDKESPPVNIIVTRE 235
QY 241 DNKDSKAIQDFVKAYQTDVEVAEAKKQFQGVYKGM 276
DB 236 DNKDAENVKKFFVQAYQSDVEVEAANKVFNGAVKGM 271

RESULT 13
ABU27902
ID ABU27902 standard; protein; 271 AA.
XX
AC ABU27902;
XX
DE 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #13429.
XX
OS Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Enterobacter cloacae.
XX
FN WO20027183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WC-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR MPI: 2003-029926/02.
DR N-PSDB; ACA31772.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening

```

PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 55826; 1766bp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 271 AA;
Query Match 56.4%; Score 780.5; DB 6; Length 271;
Best Local Similarity 56.7%; Pred. No. 9.3e-64;
Matches 152; Conservative 43; Mismatches 62; Indels 11; Gaps 3;
OY 11 ALASGIALAGCSNOSNEPAISKTAQTIKGVWAGPEQVAEVAQVAKKXNLTVEVY 70
DB 13 ALGSLALVCGCGDEKDP-----NHIKGVIVGAEQVAEVAQVAKKXNLTVEVY 64
OY 71 EFNDDYAMNSAVSKGELDANAMOHKPYLEKDSQEKLNLVIGNTFYVPLAGYSTIKXT 130
DB 65 TFDNDYVLENEALSKGDDANAFQHKPYLDDQIKDRGY-KLVGVGNTFYVPIAGYSKIKS 123
OY 131 LNELEKQATIAVNDPSNIALRALILEKQGLIKLKXNTLFFSTLIDIVENPKLVIXEVD 190
DB 124 LDELQPSQVAVENDPNTGRSLTLQKQGLILKDGVGGLPVLVDVTENPKLVIXEVL 183
OY 191 TSVAARALID--VDLAVVNNVAGOVGLTASENGVEVEKDSPEYVNIIVARADNDSKAI 248
DB 184 APQLPRSLDQAIALAVINTTASQIGLPKAKGIVFEKDSGYVNLITREDDNDAAENV 243
OY 249 QDFVAKAYQDVEAEAKKQFKQGVITKGM 276
DB 244 KKFVQAYQSDVEYQEAANKVENGAVKGM 271
RESULT 14
ABU31676
ID ABU31676 standard; protein; 271 AA.
XX
AC ABU31676;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by prokaryotic essential gene #17203.

XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Klebsiella pneumoniae.
XX
PN WO20027183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
XX
PR 05-SEP-2001; 2001US-00948993.
XX
PR 25-OCT-2001; 2001US-0342923F.
XX
PR 08-FEB-2002; 2002US-00072851.
XX
PR 06-MAR-2002; 2002US-0362699F.
XX
PA (ELITR) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Onlsen KL, Zyskind JW,
PI Wall D, Traxick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR MPI; 2003-029926/02.
XX
DR N-PSDB; ACA35546.
XX
PS Claim 25; SEQ ID NO 59600; 1766bp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 271 AA;
Query Match 56.1%; Score 775.5; DB 6; Length 271;
Best Local Similarity 55.2%; Pred. No. 2.7e-63;
Matches 153; Conservative 45; Mismatches 68; Indels 11; Gaps 3;
OY 2 NFGKINGICALAGSIALAGCSNOSNEPAISKTAQTIKGVWAGPEQVAEVAQVAKK 61
DB 4 NFKTFPAVAGLIGSLALVCGCGDEKDP-----NHIKGVIVGAEQVAEVAQVAKK 55

CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 271 AA;

Query Match 54.8%; Score 757.5; DB 6; Length 271;
 Best Local Similarity 53.2%; Pred. No. 1.3e-61;
 Matches 148; Conservative 48; Mismatches 71; Indels 11; Gaps 3;

QY 1 MNEFGKINGICALAGSGLAGCSNOSNEPAISKTAQTIKVGMAGPEQVAEVAQVAK 60
 Db 3 LKFKSIALVIALGLTLVGCGETEKAP-----NHIKVGIYVABEQVAVAKVAK 54
 QY 61 EKXNLVTELVFENDYAPNSAVSKGELDANAMOHKRYLLEDSOEGKLNLYVGNFTVP 120
 Db 55 EKXGLDELVTENDYVLPNALSKEGIDILNAFQHKRYLDOQIKDRGY-KLVSVGNSFYVP 113
 QY 121 LAGYSTKIKTLNELKDGATIAVNDPSNLARALILIEKQGLIKLKNNTNLFSTTLDIVEN 180
 Db 114 IAGYSKIKSLDELQPSQVALNDPNLGRSLILQSVGLIKKGVGLPVLVDVAVEN 173
 QY 181 PKKXIVKVDTSVAARAID--VDLAVNNNVAQVGLTSANGVFVEDKSPYVNIYA 238
 Db 174 PKXIKVTELRAPQLPSLDOQIALAIINTTVASQIGLPAKGLFVEDKESFYVNLIVA 233
 QY 239 RADNKSKAIDPFYKAYQTEVEAEAKKQFKDGVIGW 276
 Db 234 REDNKAENYKFKVQAIQSDDEVTDANKAPNGAVAKW 271

RESULT 17

ABU40467 standard; protein; 271 AA.

AC ABU40467;
 DT 19-JUN-2003 (first entry)
 DE Protein encoded by Prokaryotic essential gene #25994.
 XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
 OS *Proteus* sp.
 XX WO20027183-A2.
 XX

PD 03-OCT-2002.
 XX 21-MAR-2002; 2002WO-US009107.
 PF 21-MAR-2001; 2001US-00815242.
 XX 21-MAR-2001; 2001US-00948993.
 PR 06-SEP-2001; 2001US-03429239.
 PR 25-OCT-2001; 2001US-03429239.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 DR N-PSDB; ACA44337.
 DR WPI; 2003-029926/02.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.

Claim 25; SEQ ID NO 68391; 1766bp; English.

The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 271 AA;

Query Match 54.8%; Score 757.5; DB 6; Length 271;
 Best Local Similarity 52.7%; Pred. No. 1.3e-61;
 Matches 147; Conservative 49; Mismatches 70; Indels 13; Gaps 4;

QY 1 MNEFGKINGICALAGSGLAGCSNOSNEPAISKTAQTIKVGMAGPEQVAEVAQVAK 60
 Db 3 LKFKSIALVIALGLTLVGCGETEKAP-----NHIRVGVISSGEQVAVAKVAK 54
 QY 61 EKXNLVTELVFENDYAPNSAVSKGELDANAMOHKRYLLEDSOEGKLN-NLYVGNFTVP 119
 Db 55 EKXGLDELVTENDYVLPNALSKEGIDILNAFQHKRYL--DQIKDRNRYITVAGNFTY 112
 QY 120 PLAGYSTKIKTLNELKDGATIAVNDPSNLARALILIEKQGLIKLKNNTNLFSTTLDIVE 179
 Db 113 PLAGYSKIKSLDELQPSQVALNDPNLGRSLILIEKQGLVGLKGVGLPVLVDVAVEN 172

PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
XX Example 3; SEQ ID NO 11656; 511bp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes
CC themselves and the encoded proteins. The prokaryotes used are *Escherichia*
CC coli, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,
CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence represents an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 240 AA:
Query Match 53.8%; Score 743.5; DB 4; Length 240;
Best Local Similarity 60.0%; Pred. No. 2.1e-60;
Matches 144; Conservative 42; Mismatches 51; Indels 3; Gaps 2;
QY IKVGMAGPEQAVNEVAGOVAKKNTLVEEFNDYAMNSAVSKGELDANMOKHPYL 98
DB 2 IKGVIGABEQVAEVAQKAKKKGDELVELTNDVLTNEALSKGIDIVNMFQKPYL 61
QY 99 EKDSQKGLNNLVIVGTFPVYPLAGYSTKTKTNELKDGATIAVNDPSNLALILILEK 158
DB 62 DQQIKDNGY-KLVAVGVTFFVPIAGYSKTKSLDELQPSQIAVPNDPTNLGRSLILLOQ 120
QY 159 QGLTKLMDNTMLSTTLDIYENPKKVIKEDVTSVAARA-DD--VDLAVNNNYAGQVGL 216
DB 121 VGLTKLMDGVGLTFTSLDIYENPKKIVELKAPQLPRSLDDAIALAVINTTYAGQIGL 180
QY 217 TASENGYFVEDKDSPPYNNIIVARADNKSKAIDPFVAYQDTEVEAKKQFQDGYIKGM 276
DB 181 TPADGIFVEGKESFPYNNLIIVARADNKAENVKKFQVAYQSDVEYEAANKI FNGAVKGM 240

RESULT 20
ABU19780
ID ABU19780 standard; protein; 295 AA.
XX
XX ABU19780;
AC
XX 19-JUN-2003 (first entry)
DT
XX
DE Protein encoded by Prokaryotic essential gene #5307.
DE
XX
KM Antisense; prokaryotic essential gene; cell proliferation; drug design.
KM
OS *Borrelia cepacia*.
XX
XX WO200277183-A2.
XX
PD 03-OCT-2002.
PF
XX 21-MAR-2002; 2002WO-US009107.
PF
XX 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948593.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX

PA (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohleen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI; 2003-029926/02.
DR
DR N-PSDB: ACA23650.
XX
PS Claim 25; SEQ ID NO 47704; 1766bp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation; (7) identifying a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 295 AA:
Query Match 40.6%; Score 562; DB 6; Length 295;
Best Local Similarity 44.5%; Pred. No. 1.9e-43;
Matches 118; Conservative 47; Mismatches 88; Indels 12; Gaps 3;
QY 12 LASGIALAGCSNQSNEPAISKTAOTKVGVMAGPEQAVNEVAGOVAKKNTLVEE 71
DB 37 LATGAALTLKSAHED-----KVIKGTAVAGPSEFVQVQKAKKEGGLNVAVIE 88
QY 72 FNDYAMPNSAVSKGELDANMOKHPYLEKDSQKGLNNLVIVGTFPVYPLAGYSTKTKTL 131
DB 89 FNDYQPMNALDSGLDANSFOHOPYLDSQVCKRGY-KIVSAGITLISPIGVSKKPKSL 147
QY 132 NELKDGATIAVNDPSNLALILILEKQGLTKLMDNTML--FSTTLDIYENPKKVIK 188
DB 148 KDLPGAKLAVPNDSENRRALLLLOQGVTKLAKAGCTGANNATVDIAENPKKLISE 207
QY 189 VDTSVARAIDVDLAVNNNYAGQVGLTASENGYFVEDKDSPPYNNIIVARADNKSKA 248
DB 208 LDAQULPRVLSDVDAVINTTYALANLQPTKDAIALESITSPYANNIIVARAKKDQPVW 267
QY 249 QDFVAYQDTEVEAKKQFQDGYI 273
DB 268 KDLVAYQSDVEYEAANKI FNGAVKGM 292

RESULT 21

ABU23447
ID ABU23447 standard; protein: 262 AA.
XX
XX ABU23447;
AC
XX 19-JUN-2003 (first entry)
DT
XX Protein encoded by Prokaryotic essential gene #8974.
DE
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
XX Bordetella pertussis.
OS
XX W0200277183-A2.
PN
XX 03-OCT-2002.
PD
XX 21-MAR-2002; 2002WO-US009107.
PF
XX 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
DR N-PSDB; ACA27317.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 51371; 1766bp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed; (12) determining the extent
XX to which each of the strains is present in a culture or collection of
XX strains; or (13) identifying the target of a compound that inhibits the
XX proliferation of an organism. The antisense nucleic acids are useful for
XX identifying proteins or screening for homologous nucleic acids required
XX for cellular proliferation to isolate candidate molecules for rational
XX drug discovery programs, or for screening homologous nucleic acids
XX required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
XX *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
XX the target prokaryotic essential genes. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pat_sequences
XX
XX Sequence 262 AA;

Query Match 39.8%; Score 551; DB 6; Length 262;

Best Local Similarity 43.1%; Pred. No. 1,66-42; Matches 118; Conservative 52; Mismatches 86; Indels 18; Gaps 5;

QY 1 MNFGKINGICALASGIALAGCSNQSNBPAISKTAQTIKYGVWAGPQVAEVA-GQVA 59
DB 3 MNFVR---SALASFLLAGAAQ-----AEKLVGATGVFAEILEVKKPALA 48
QY 60 KEKNLITVELVFENYAMPNSAVSGEDDANAMQKPTLEKDSQSKGLNNIYIGENTFY 119
DB 49 KGVGLDIKIV--FTTYVQPNQLADKQDPAFFQHPYLDFTFNKOR-KTNLVSGLVAHE 105
QY 120 PLAGYSTKIKTINLEKDGATIAVPNDPSNLARALITLEKQGLIKKDWNLFTSLDIVE 179
DB 106 PFGYSKTKISLAELKDGATIALPNDPSNSGALLLLKQGLKKDKDPSNIYATIDIAE 165
QY 180 NPKULVKEVDTSVAPRAIDVDIAVNNNTAGQVGLTASENGVEDEKDSFYNIIVAR 239
DB 166 NPKULKEFLELAAMPDRSFDDLDLALINTVYALAEAGLPTPRDALFIEGADSEYANLVAAR 225
QY 240 ADNKSKAIDQDFKAYQGTDEVEAEAKKQKQVI 273
DB 226 PNKDPAVAKKLVNLAHSEAVKFTIEKIKAVV 259

RESULT 22

ABU22451

ID ABU22451 standard; protein: 272 AA.

AC ABU22451;

XX 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #7978.

KM Antisense; prokaryotic essential gene; cell proliferation; drug design.

OS Burkholderia mallei.

PN W0200277183-A2.

XX 03-OCT-2002.

PF 21-MAR-2002; 2002WO-US009107.

PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

PS Claim 25; SEQ ID NO 50375; 1766bp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the

CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
CC

SQ Sequence 272 AA;

Query Match 39.7%; Score 549.5; DB 6; Length 272;

Best Local Similarity 42.5%; Pred. No. 2.4e-42;

Matches 114; Conservative 52; Mismatches 91; Indels 11; Gaps 3;

QY 9 ICALASGATAGCGSQNSNEPAIAKTAQITKVGWMAPEQVAEVAQVAKKXNLTVE 68
DB 10 VSAVAAAGAAASVS-----VGAQADQVIRKVTGAVGPDQVQVYQKAKKQGLDVK 62
QY 69 LVEFNDVAMPNSAVSKGELDANAMQHKRYEKDSQKGLNNLVITGNTFVYPLAGYSTKT 128
DB 63 VLEFNDVYQPAALADSDGLDANSFQHQPYLDSQYKQSGY-KIVASGLTYSPIGVSKKF 121
QY 129 KTLWEKQATIVAPNDPSNLARALILIEKQGLIKLKDNTNL-----FSTTLDIYENPKKV 135
DB 122 KSLWELEPAKAVVALPNDPSNENRALLLQIQGVTKLAKAGAGTGSNAIVLDVLENPKKLK 131
QY 186 IKEYDTSVAARAIIDVDLAVVNNYAGVGLTASENGVEEDKDSFYVNIIVARADKDS 245
DB 182 LTELDAAGLPRVLSVDVAIVNTVYALANLQPKDAIALESLTSPYANLIAVARAKDQ 241
QY 246 KAIQDFVAKYQTDVEAEAKQKQFDGVI 273
DB 242 FVWKGLVYKAYOSPVEKFEITQFQKGSV 269

RESULT 23

ABU33203

ID ABU33203 standard; protein; 256 AA.

AC ABU33203;

DT 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #18730.

KM Antisense; prokaryotic essential gene; cell proliferation; drug design.

OS Legionella pneumophila.

XX WO200277183-A2.

PD 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 06-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
PA Wang L, Zmudis C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall L, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
DR N-PSDB; ACA37073.

XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 61127; 1766bp; English.

CC The invention relates to an isolated nucleic acid comprising any one of
CC the 621 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
CC

SQ Sequence 256 AA;

Query Match 39.4%; Score 545; DB 6; Length 256;

Best Local Similarity 45.7%; Pred. No. 5.7e-42;

Matches 118; Conservative 37; Mismatches 91; Indels 12; Gaps 3;

QY 16 IALAGCSQNSNEPAIAKTAQITKVGWMAPEQVAEVAQVAKKXNLTVELEFNDY 75
DB 11 ISLVACS-----SKPAPVTLVIGTIGAGETELTETAKOVAKEKGLIKIIEFNDY 61
QY 76 AMPNSAVSKGELDANAMQHKRYEKDSQKGLNNLVITGNTFVYPLAGYSTKTINLEK 135
DB 62 NLPNEALDQSDADAVYQHLPLKRAILSHG-DIQALGRFVYPMGITSKXNTLSLP 120
QY 136 DGATIAVNDPSNLARALILIEKQGLIKLKDNTNLFTTLDIYENPKKVLEKEYDTSVAA 195
DB 121 ENGIIAIVNDPENEMRAFLLEKXHLITLKNTTN--SGIQDIESNPQKFKREIDAQLP 178
QY 196 PAIDVDLAVVNNYAGVGLTASNGVFEEDKDSFYVNIIVARADNDKSAIQDFVAKY 255
DB 179 RVLPPVDVAIVNTTLPALPAGLSPSKDALFTGKDSFYVNIIVIRDRTEXPQLLEFVAL 238

QY 256 QTDVEAEAKKQFQDGV 273
DB 239 NSXEVEKAXNLFQEDAI 256

RESULT 24

AB021332
ID AB021332 standard; protein; 268 AA.
XX
XX AB021332;
XX
XX 19-JUN-2003 (first entry)
XX
XX Protein encoded by Prokaryotic essential gene #6859.
XX
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
OS Burkholderia fungorum.
XX
XX MO200277183-A2.
PD
PD 03-OCT-2002.
XX
XX 21-MAR-2002; 2002MO-US009107.
PF
XX
XX 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
PI Wall D, Tirawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
XX
XX N-PSDB; ACA25202.
DR
XX
XX Claim 25; SEQ ID NO 49256; 1766pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the farthest orthologous essential genes. Note: The sequence data for this

CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX
SQ Sequence 268 AA;

Query Match 39.0%; Score 540; DB 6; Length 268;
Best Local Similarity 44.6%; Pred. No. 1.8e-41;
Matches 108; Conservative 49; Mismatches 83; Indels 2; Gaps 2;

QY 35 AAQTIKGVMAQPBQAEVAAGVAKKQNTLVLEFNDYAMNSAASKGELDANAMQH 94
DB 29 AEDTIKVGVTGSPHAQIMEVKTVA-AKNGLNIVKESDIYQFPAALAGSLDANSYQH 87
QY 95 KPYLEKDSQEKGLNNLVVGNTPFVPLAGYSTKIKITNELKDGATTIAPNDPSMLARALI 154
DB 88 DPLYGAQYKDRGY-KLIRVADTVTYPMGIYSKVKSLAEILOGAKIAPNDPTNGRALL 146
QY 155 LLEKQGLIKLDNNLTSTTDIYENPKLVYKESVDSVAPARAIDDDVDLAVNNNYAGQV 214
DB 147 LLOKQGLIKLRADAGLKTATPDIYDNPCKLIVELDAAQIPRSINDVDAAAINTFAMEA 206
QY 215 GLTASENGVFVEDKDSFVYNIIVARADKDSKAIODFVKAQTDVEAEAKKQFQDGV 274
DB 207 GLKPKQDAIATEDPGPFYVNIATREADRNKFWAKLVAAVHSPEVKKFVSKEGSAVIT 266
QY 275 GW 276
DB 267 AW 268

RESULT 25

AB021415
ID AB021415 standard; protein; 270 AA.
XX
XX AB021415;
AC
XX
XX 19-JUN-2003 (first entry)
DT
XX
XX Protein encoded by Prokaryotic essential gene #6942.
DE
XX
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
KW
XX
XX Burkholderia fungorum.
OS
XX
XX MO200277183-A2.
PN
XX
XX 03-OCT-2002.
PD
XX
XX 21-MAR-2002; 2002MO-US009107.
PF

XX
XX 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
PI Wall D, Tirawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
XX
XX N-PSDB; ACA25285.
DR
XX
XX WPI; 2003-025926/02.
PR
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
PS Claim 25; SEQ ID NO 49339; 1766pp; English.
XX
XX
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression

QY 189 VDTSVARAIDVDVLA VNNVAGVGLTASENGVEVEDKDSPPYNIIVARADNKSRAI 248
DB 178 LDAAGQPRSLDDTDA SAVNTNFPALFAGLDPSKALVRESAESPYANVLVREDDKDRADL 237
QY 249 QDFVYAYQTDVEVEAEAKKQFKDGVIKGW 276
DB 238 RKLVSIVQSAFVREFLIGTKYKGA VAAW 265

RESULT 27
ABB53625
ID ABB53625 standard; protein; 286 AA.
AC ABB53625;
XX
XX 29-AUG-2003 (revised)
DT 16-MAY-2002 (first entry)
XX
XX Lactococcus lactis protein plpB.
DE Lactococcus lactis protein plpB.
XX
XX Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
XX
XX Lactococcus lactis; IL1403.
OS
XX
XX FR2807446-A1.
PN
XX
XX 12-OCT-2001.
PD
XX
XX 11-APR-2000; 2000FR-00004630.
PF
XX
XX 11-APR-2000; 2000FR-00004630.
PR
XX
XX (INRG) INRA INST NAT RECH AGRONOMIQUE.
PA
XX
XX Bolotine A, Sorokine A, Renault P, Ehrlich SD;
PI
XX
XX WPI; 2002-043418/06.
DR
XX
XX New nucleotide sequence useful in the identification or Lactococcus
PT lactis and related species.
XX
XX
XX Claim 6; SEQ ID NO 327; 2504bp; French.

CC The present invention is related to a Lactococcus lactis nucleotide
CC sequence (AA090521) and related proteins (ABB53300-ABB55621). The nucleic
CC acid sequence is useful in the detection and/or amplification of nucleic
CC acid sequence, particularly to identify Lactococcus lactis or related
CC species. The proteins of the invention are useful for the biosynthesis or
CC biodegradation of a composition of interest. The invention helps research
CC in lactic bacteria, particularly useful in the production of yogurt and
CC cheese. Note: The sequence data for this patent is based on equivalent
CC patent WO200177334 (published 18-OCT-2001) which is available in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences. (Updated on 29-AUG-2003 to
CC standardise OS field)
XX
XX
XX Sequence 286 AA:

Query Match 36.8%; Score 509.5; DB 5; Length 286;
Best Local Similarity 40.5%; Pred. No. 1.3e-38;
Matches 111; Conservative 58; Mismatches 90; Indels 15; Gaps 8;

QY 9 ICALGAGIALAGSNGSNEPAISKTAOTIKVWAGSEQ--AAVEVAGVAKEXYNT 66
DB 17 IVAIAAFIG-GQKSGAN-----KTVNKTAKIGITGTEKEDDSIQATSKAKKXGTT 69
QY 67 VELVEFNDAVAMPNSAVSKGELDANAMQHKPYEKSDCEKGLNNLVIVNTFVYPLAGYST 126
DB 70 LKFTHTFTDYQPTALKNQGDIDINAFQHYAFL-KAMNKANNGVLVAIGTVISPISVSK 128
QY 127 KITLNEIDCGATIAVPNDPSNARALLILEKGLFKLKDNTNLFSTTLDIVENPKXIYI 186

DB 129 QLNKISDIKEGATTAVPNDASNESRALVYLKAGLIKLDVSGOTLATVKTITSPKXNLI 188
QY 187 KEVDPSVARAIDVDVLA VNNVAGVGLTASENGVEVE--DDKS--PYNIIIVARADNK 243
DB 189 KELDASQIARALDSVDAAVNNNNAVTRAGLKSD-AIFTEFVNKDSQCMIIIVANKDE 247
QY 244 DSKAIQDFVYAYQTDVEVEAEAKKQFKD-GVIKGW 276
DB 248 NNTVYKDVYKAYETEATKKTIAKAYPKSTIPAW 281

RESULT 28
ABU39768
ID ABU39768 standard; protein; 261 AA.
AC ABU39768;
XX
XX 19-JUN-2003 (first entry)
DT
XX
XX Protein encoded by prokaryotic essential gene #25295.
DE
XX
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
XX
XX Pseudomonas putida.
OS
XX
XX WO200277183-A2.
PN
XX
XX 03-OCT-2002.
PD
XX
XX 21-MAR-2002; 2002WO-US009107.
PF
XX
XX 21-MAR-2001; 2001US-00815242.
PR
XX
XX 06-SEP-2001; 2001US-00948993.
PR
XX
XX 25-OCT-2001; 2001US-0342923P.
PR
XX
XX 08-FEB-2002; 2002US-00072851.
PR
XX
XX 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-). ELITRA PHARM INC.
PA
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen XL, Zyskind JW;
PI
XX
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Foreyth RA, Xu HH;
XX
XX WPI; 2003-023926/02.
DR
XX
XX N-PEDB; ACA43638.

CC New antisense nucleic acids, useful for identifying proteins or screening
CC for homologous nucleic acids required for cellular proliferation to
CC isolate candidate molecules for rational drug discovery programs.
XX
XX
XX Claim 25; SEQ ID NO 67692; 1766bp; English.

CC The invention relates to an isolated nucleic acid comprising any one of
CC the 623 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target in a culture that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for

identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 261 AA;

Query Match 36.6%; Score 506.5; DB 6; Length 261;
Best Local Similarity 41.3%; Pred. No. 2,2e-38;
Matches 109; Conservative 52; Mismatches 88; Indels 15; Gaps 4;

11 ALASGILAGCSNQSNNPAISKTAQTIKVGWAGBEQAVAA-VAGQVAKKYNLTVEL 69
9 ALAAISFSGIA-----AAEKLVAAATPVPFAHEILIFETLAKGEVDLQIV 57
QY 70 VEENDVAMPNSAVSKGELDANAMQHKPYLEKDSQEKGLNNLVTVGNTFPYPLAGYSTKIK 129
DB 58 --FDYQVPNVQYDQKRLDANYFQTLPYLQNFNEGKTHLETYVG-VHVEPFGYSKKYK 114
QY 130 TINEIKGATITAVPNDNSNARALILIEKGLIKLKDNNLFTSLDIYENPKLVIKEV 189
DB 115 ALSELKGAIVAAIPNEGSGRALLLQKAGLITLKDPPKVALTPKDIENPKLKEFEL 174
QY 190 DTSAARAIDVDLAVANNVYAGVGLTASENGVEYEDKSPYNIIVARADNKSKAIO 249
DB 175 ESAMLPRLVDQVLDLMTNTYALAEGLNPAKDALVEGADSPVNLVLARPDKSEALQ 234
QY 250 DFKAVYDTDEVEAEAKQKQFDGYI 273
DB 235 KLAQALTSPEVKAFAKCYGAVL 258

RESULT 29

AAU36517 standard; protein; 260 AA.

AAU36517;
14-FEB-2002 (first entry)
Pseudomonas aeruginosa cellular proliferation protein #507.
Antisense; prokaryotic cellular proliferation protein; antibiotic;
antibacterial; drug design.
Pseudomonas aeruginosa.
WO200170955-A2.
27-SEP-2001.
21-MAR-2001; 2001WO-US009180.
21-MAR-2000; 2000US-0191078P.
23-MAY-2000; 2000US-0206848P.
26-MAY-2000; 2000US-0207727P.
23-OCT-2000; 2000US-0242578P.
27-NOV-2000; 2000US-0253625P.
22-DEC-2000; 2000US-0257931P.
16-FEB-2001; 2001US-0269308P.
(ELIT-) ELITRA PHARM INC.
Haseelbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GU, Yamamoto RT, Xu HH;
WPI; 2001-611495/70.
N-PSDB; AAS54376.

XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
PS Example 3; SEQ ID NO 12110; 511p; English.

CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes
CC themselves and the encoded proteins. The prokaryotes used are *Escherichia*
CC *coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,
CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence represents an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 260 AA;

Query Match 36.3%; Score 502.5; DB 4; Length 260;
Best Local Similarity 42.6%; Pred. No. 5e-38;
Matches 106; Conservative 52; Mismatches 84; Indels 7; Gaps 3;

QY 29 AAIKTAQTIKVGWAGBEQAVAEVAGVY---AKEKNLTVELVEFNDAVAMPNSAVSK 84
DB 12 AALGILTAQAESTLVAAATPVPFAHEILANVYKPLAKGVDLKI--EFTYVQPNVQVSE 69
QY 85 GELDANAMQHKPYLEKDSQEKGLNNLVTVGNTFPYPLAGYSTKIKTINELKDGATIVEN 144
DB 70 KRLDANFQHPQYIDENKAKG-TDLVAVTGVHIEPLGAVSSKKKIDELPSGATVIVP 128
QY 145 DPSNLARALILIEKGLIKLKDNNLFTSLDIYENPKLVIKEVDVSAARAIDVDLA 204
DB 129 DATNGRALLLDRAGVYKDKNSITATPDKDIDVKNKIKRLEAAATPRLVITQYDMA 188
QY 205 VVNNVYAGVGLTASENGVEYEDKSPYNIIVARADNKSKAIODFVKAQYQDEVEAEA 264
DB 189 LINTNYALAEKLANTXALAIEGSDSPYNIIVARPDNKSADANQKALHSAEIKQFI 248
QY 265 KKQKFDGYI 273
DB 249 QEKYXGAVV 257

RESULT 30

ABU38920 standard; protein; 260 AA.

ABU38920;
19-JUN-2003 (first entry)
Protein encoded by Prokaryotic essential gene #24447.
Antisense; prokaryotic essential gene; cell proliferation; drug design.
Pseudomonas aeruginosa.
WO200277183-A2.
03-OCT-2002.
21-MAR-2002; 2002WO-US009107.
21-MAR-2001; 2001US-00815242.


```
Db 7 LTLASVY-----SVSASAEEKLIVATPIPAETLEIKPTLAEGVDLQI 53
Qy 68 ELVEFNDYAMPNSAVSGEILDANAMQHPYLEKDSQEKGLNNVIYGNTEFVEYLAGYSTK 127
Db 54 KV--FTDYQPNVAVAEKRLDANVFQTLPLENNKSGKGNLTVVG-VHVEPFGYSRK 110
Qy 128 IKTINELKDGATTAVPNDPSNLAPALILEKQGLIKLKDNTNLFSTLLDIVENPKKLVIR 187
Db 111 YKSLAELPDGATVAIPNEGNSGRALLLQKAGLLKPKNNALATPKDIAENPKNLKPK 170
Qy 188 EVDTSVAAARAIDVDVLAANNVAGVGLTASENGVEDEKDSFYNIIVARADNKSQA 247
Db 171 ELBSALLPRVLDQVDLDLINTNTVALEKALNPACDALVLEBRDSPYNYVVARPNKDSQA 230
Qy 248 IQDFVKAAYQTDVEVAEAKKQFKDQVI 273
Db 231 LKKLSALTSPVAVAFIEKKYAGAVV 256

RESULT 32
ABU15545
ID ABU15545 standard; protein; 259 AA.
XX
AC ABU15545;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by prokaryotic essential gene #1072.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Pseudomonas aeruginosa.
XX
PN MO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948893.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KT, Zyskind UW,
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI; 2003-029926/02.
XX
XX N-PSDB; ACA19415.
XX
XX Claim 25; SEQ ID NO 43469; 176bp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
```

```
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 259 AA;
XX
XX Query Match 36.0%; Score 497.5; DB 6; Length 259;
XX Best Local Similarity 41.4%; Pred. No. 1.5e-37;
XX Matches 110; Conservative 51; Mismatches 88; Indels 17; Gaps 4;
Qy 9 ICALAGIATLACGNSQNEPAAISKTAQTIKYGVNAGPQOAVAE-VAGQVAKKYNLTV 67
Db 7 LTLASVY-----SVSASAEEKLIVATPIPAETLEIKPTLAEGVDLQI 53
Qy 68 ELVEFNDYAMPNSAVSGEILDANAMQHPYLEKDSQEKGLNNVIYGNTEFVEYLAGYSTK 127
Db 54 KV--FTDYQPNVAVAEKRLDANVFQTLPLENNKSGKGNLTVVG-VHVEPFGYSRK 110
Qy 128 IKTINELKDGATTAVPNDPSNLAPALILEKQGLIKLKDNTNLFSTLLDIVENPKKLVIR 187
Db 111 YKSLAELPDGATVAIPNEGNSGRALLLQKAGLLKPKNNALATPKDIAENPKNLKPK 170
Qy 188 EVDTSVAAARAIDVDVLAANNVAGVGLTASENGVEDEKDSFYNIIVARADNKSQA 247
Db 171 ELBSALLPRVLDQVDLDLINTNTVALEKALNPACDALVLEBRDSPYNYVVARPNKDSQA 230
Qy 248 IQDFVKAAYQTDVEVAEAKKQFKDQVI 273
Db 231 LKKLSALTSPVAVAFIEKKYAGAVV 256

RESULT 33
ABU25052
ID ABU25052 standard; protein; 263 AA.
XX
AC ABU25052;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by prokaryotic essential gene #10579.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Clostridium difficile.
XX
PN MO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948893.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
```

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
DR N-PDSB; ACA28922.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 52976; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation; (7) identifying a gene in an operon required for
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 263 AA;

Query Match 35.7%; Score 494; DB 6; Length 263;
Best Local Similarity 41.7%; Pred. No. 3.1e-37;
Matches 111; Conservative 50; Mismatches 87; Indels 18; Gaps 6;

11 ATASGIALA--GCSNOSNEPAISKTAOTIKGVAGAGPEQVAEYAGVAREK--YNTLV 67
DB ALVSAIAISAVGCSNED-----KKILVGSNSNHAKILEVAKXLEKGGIDLEV 59
QY 68 ELVEFNDYAMPNSAVSKGELDANANOHKRYLEKDSOEKLANLVVGTFTVPLAGYSTK 127
DB KI--PDVYALPNVADLEGSLDANFPHIIFLEBETVEYKEGY--KLTYSKTHIEMPFYSSEK 116
QY 128 IKTNELKKGATIAVNDPSNLARALITLEKGLIKLKNNTLVFSTTLTIVENPKKLVTK 187
DB VTSJDLKKGAVIAVNDATNGARALKLAKKKLLEVGKGLI--TKKDITNPKXIVTK 174
QY 188 EVDTSVAARADVDLIAVNNNVAQGVLTASENGVFVYDKDSPYVNIIVARADKSKA 247
DB ENNAEQLPTVLKXVDGAVINSNALTANINPTDAIVIESSDSPYVNIIVARADKSKA 234
QY 248 IODFVAKYOTDEVEAEAKKOFKGV 273
DB IKALSDAMNSKEVKKFIQDEYKGSIV 260

RESULT 34

ABU39703
ID ABU39703 standard; protein; 256 AA.
XX
AC ABU39703;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #25730.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS *Pseudomonas putida*.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002MO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
XX N-PDSB; ACA43573.
DR
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 67627; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation; (7) identifying a gene in an operon required for
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 256 AA;

[illegible]


```
PR 16-MAY-1997; 97US-004655P.  
PR 14-NOV-1997; 97US-0066009P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Kunsch CA, Choi GH, Bailey C, Hromockyj A;  
XX WPI; 1999-070095/06.  
XX N-PSDB; AAX20014.  
XX  
XX New isolated Enterococcus faecalis polynucleotides - used to develop  
PT products for the detection of Enterococcus and for use in vaccines for  
PT prevention or attenuation of Enterococcus infection.  
XX  
XX Claim 9; Page 86; 301pp; English.  
XX  
XX The present sequence represents a protein isolated from Enterococcus  
CC faecalis. The present invention describes genes, proteins and antigenic  
CC polypeptides isolated from E. faecalis. The proteins can be used in  
CC vaccines for preventing or attenuating an infection caused by a member of  
CC the Enterococcus genus in an animal. They can also be used for detecting  
CC Enterococcus antibodies in a sample. The nucleotide sequences can be used  
CC for detecting Enterococcus nucleic acids. Products from the present  
CC invention can also be used for screening compounds to identify agonists  
CC and antagonists of E. faecalis protein activity  
XX  
XX Sequence 277 AA;  
SQ  
  
Query Match 34.8%; Score 481; DB 2; Length 277;  
Best Local Similarity 36.8%; Pred. No. 5.4e-36;  
Matches 102; Conservative 62; Mismatches 101; Indels 12; Gaps 5;  
  
QY 3 FGIKINGICATLALAGCSNQSNEPAISKTAQTIKGVMAAPBQAVAEVAGQVAKK 62  
DB 4 FSKLIGLIGVLA-FTIAGCAGS-----VKDTKETVKGAVGTGKNDEWESVDRKKK- 56  
QY 63 YNLTVLEVFENDYAMPNSAVSGKGLDANAMQHKPYLEKSOEKGANNLVYGTFTVYPLA 122  
DB 57 -NIDQLVLFDTYQPAALAEKEIDLNAPHQIFLDNKNKEHG-TKLVSIGNTVNAAPLG 114  
QY 123 GYSTKIKTLNELKDGATIAVPNDPSNLARALILIEKQGLIKLKDNTLFTSLDIVENPK 182  
DB 115 IYANKLKDITKIKDGEIAPNDPTNGRALIILQTAGLILKVDPAKQQLPTVSDITENKR 174  
QY 183 KLVKEVDTSVARALDDVDLAVNNNNYAGVGLTASENGVFE--DKDSPYNIIVAR 239  
DB 175 QLKTELDATITATRALQVDASVINSGMADVAGTTPDKAIFLEPVEKAKPYNIIVAR 234  
QY 240 ADNKSRAIQDFVAKYQTDVEVEAKKQFQDGVTKGW 276  
DB 235 EEDQENKLYQKVVEYEQEETKVIATETSKGANVPAM 271  
  
RESULT 37  
ABP43243  
ID ABP43243 standard; protein; 277 AA.  
XX  
XX ABP43243;  
XX  
XX 05-AUG-2002 (first entry)  
XX  
XX E faecalis EF009 protein.  
XX  
XX Enterococcus; vaccine; gastrointestinal disease; diagnosis; antibiotic.  
XX  
XX Enterococcus faecalis.  
XX  
XX US2002045737-A1.  
XX  
XX 18-APR-2002.  
XX  
XX 04-MAY-1998; 98US-00071035.  
XX
```

```
PR 04-MAY-1998; 98US-00071035.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Choi GH, Bailey C, Hromockyj A, Kunsch CA;  
XX WPI; 2002-425450/45.  
XX N-PSDB; ABN97999.  
XX  
XX New genes and polypeptides from Enterococcus faecalis, useful as vaccines  
PT for preventing, treating or attenuating an infection caused by a member  
PT of the Enterococcus genus in an animal, particularly E. faecalis.  
XX  
XX Claim 9; Page 35; 255pp; English.  
XX  
XX The present invention provides the protein and coding sequences of a  
CC number of polypeptides from Enterococcus faecalis. The proteins can be  
CC used as vaccines for preventing or attenuating an infection caused by a  
CC member of the Enterococcus genus in an animal, particularly E. faecalis.  
CC The polynucleotide is also useful for preventing or treating E. faecalis  
CC infection. The present sequence is a protein of the invention  
XX  
XX Sequence 277 AA;  
SQ  
  
Query Match 34.8%; Score 481; DB 5; Length 277;  
Best Local Similarity 36.8%; Pred. No. 5.4e-36;  
Matches 102; Conservative 62; Mismatches 101; Indels 12; Gaps 5;  
  
QY 3 FGIKINGICATLALAGCSNQSNEPAISKTAQTIKGVMAAPBQAVAEVAGQVAKK 62  
DB 4 FSKLIGLIGVLA-FTIAGCAGS-----VKDTKETVKGAVGTGKNDEWESVDRKKK- 56  
QY 63 YNLTVLEVFENDYAMPNSAVSGKGLDANAMQHKPYLEKSOEKGANNLVYGTFTVYPLA 122  
DB 57 -NIDQLVLFDTYQPAALAEKEIDLNAPHQIFLDNKNKEHG-TKLVSIGNTVNAAPLG 114  
QY 123 GYSTKIKTLNELKDGATIAVPNDPSNLARALILIEKQGLIKLKDNTLFTSLDIVENPK 182  
DB 115 IYANKLKDITKIKDGEIAPNDPTNGRALIILQTAGLILKVDPAKQQLPTVSDITENKR 174  
QY 183 KLVKEVDTSVARALDDVDLAVNNNNYAGVGLTASENGVFE--DKDSPYNIIVAR 239  
DB 175 QLKTELDATITATRALQVDASVINSGMADVAGTTPDKAIFLEPVEKAKPYNIIVAR 234  
QY 240 ADNKSRAIQDFVAKYQTDVEVEAKKQFQDGVTKGW 276  
DB 235 EEDQENKLYQKVVEYEQEETKVIATETSKGANVPAM 271  
  
RESULT 38  
ABU88271  
ID ABU88271 standard; protein; 277 AA.  
XX  
XX ABU88271;  
XX  
XX 07-JUL-2003 (first entry)  
XX  
XX E. faecalis novel protein #15.  
XX  
XX Vaccine; endocarditis; bacteraemia; urinary tract infection; UTI;  
XX intraabdominal infection; soft tissue infection; neonatal sepsis.  
XX  
XX Enterococcus faecalis.  
XX  
XX US2003017495-A1.  
XX  
XX 23-JAN-2003.  
XX  
XX 29-JUL-2002; 2002US-00206576.  
XX  
XX 06-MAY-1997; 97US-0044031P.  
XX 16-MAY-1997; 97US-004655P.  
XX 14-NOV-1997; 97US-0066009P.  
XX
```

RESULT 39
 ABU29285
 ID ABU29285 standard; protein; 277 AA.
 AC ABU29285;
 DT 19-JUN-2003 (first entry)
 DE Protein encoded by Prokaryotic essential gene #14812.
 KM Antisense; prokaryotic essential gene; cell proliferation; drug design.
 OS Enterococcus faecalis.
 SN WO200277183-A2.
 XX
 XX
 XX 03-OCT-2002.
 PD
 PD 21-MAR-2002; 2002WO-US009107.
 PF
 PF 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 23-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Walil D, Treawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 PI WPI; 2003-029926/02.
 DR N-PSDB; ACA33155.
 XX
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 57209; 1766bp; English.
 XX
 XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 613 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pat_sequences

XX Sequence 277 AA;
SQ

Query Match 34.8%; Score 481; DB 6; Length 277;
Best Local Similarity 36.8%; Pred. No. 5.4e-36;
Matches 102; Conservative 62; Mismatches 101; Indels 12; Gaps 5;

QY 3 FKGINGICALASGIALAGCSNOSNEPAISKTAQTIKVGWAGPEQVAEVAQVAKK 62
DB 4 FSKLIGLIGVLA-FTIAGCAGS-----VKDTKETVKGVGVTXNDWESVQDLKKK- 56
QY 63 YNLTVLVEFNDYAMPNSAVSKGELDANAMQHKPYLEKDSQEKGLNLTIVGNTFVYPLA 122
DB 57 -NIDQLVEFTDYTPNPAALAEKEIDLNAFQHFIDNVRKEHG-TKLVSIGNTVNAFLG 114
QY 123 GYSTRIKTLNELKDGATIAVNDPSNLARALILEKQGLIKLKONTNLFSTLIDIVENPK 182
DB 115 IYANKLKQITIKIDGGEIAPNDPTNGRALIILQTAGLIKVDPAKQQLPTVSDITENKR 174
QY 183 KLVYKEVDTVAARAIDVDLAVVNNNYAGQVGLTASENGVEFVE---DKDSPYNIIVAR 239
DB 175 QLKITELDATQTAARALQVDVASVNSGAVDAGYTPDKDAIFLEPVNEKAKPYNIIVAR 234
QY 240 ADNKSRAIDQFVKAYOTDEVEAEAKKQFQGVIKGW 276
DB 235 EBDQENKLYQKVVEEYQOEETKVIATETSKGANVPAM 271

RESULT 40

ABU13522
ID ABU13522 standard; protein; 277 AA.

AC ABU13522;

XX 26-FEB-2003 (first entry)

DE Enterococcus faecalis EF040 polypeptide #15.

XX EF040; immunostimulant; antibacterial; gene mapping.

OS Enterococcus faecalis.

XX US6448043-B1.

XX 10-SEP-2002.

PF 04-MAY-1998; 98US-00071035.

PR 06-MAY-1997; 97US-0044031P.

PR 16-MAY-1997; 97US-0046655P.

PR 14-NOV-1997; 97US-0066009P.

PR 14-NOV-1997; 97US-0066099P.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Choi GH, Bailey C, Hromockyj A, Kunach CA;

DR WPI; 2003-089120/08.

DR N-PSDB; ABX61569.

XX Example 1; Col 55-56; 146pp; English.

XX The invention relates to polynucleotide fragments of a gene from
CC Enterococcus faecalis, EF040, and the polypeptides encoded by them. The
CC polypeptides are useful in detecting E. faecalis, as epitope tags, as
CC molecular weight markers on SDS-PAGE gels or for molecular sieve gel
CC filtration columns, in generating antibodies that specifically bind to
CC the E. faecalis polypeptides, in generating an immune response against E.
CC faecalis and other Enterococcus species and as vaccines against other

CC bacterial genera. The polynucleotides are useful as probes for gene
CC mapping and for identifying E. faecalis in biological samples. Sequences
CC ABU13508-ABU13755 represent EF040 polypeptides of the invention. Note:
CC The sequence data for this patent can also be obtained from USPRO at
CC seqdata.uspro.gov/sequence.html

XX Sequence 277 AA;
SQ

Query Match 34.8%; Score 481; DB 6; Length 277;
Best Local Similarity 36.8%; Pred. No. 5.4e-36;
Matches 102; Conservative 62; Mismatches 101; Indels 12; Gaps 5;

QY 3 FKGINGICALASGIALAGCSNOSNEPAISKTAQTIKVGWAGPEQVAEVAQVAKK 62
DB 4 FSKLIGLIGVLA-FTIAGCAGS-----VKDTKETVKGVGVTXNDWESVQDLKKK- 56
QY 63 YNLTVLVEFNDYAMPNSAVSKGELDANAMQHKPYLEKDSQEKGLNLTIVGNTFVYPLA 122
DB 57 -NIDQLVEFTDYTPNPAALAEKEIDLNAFQHFIDNVRKEHG-TKLVSIGNTVNAFLG 114
QY 123 GYSTRIKTLNELKDGATIAVNDPSNLARALILEKQGLIKLKONTNLFSTLIDIVENPK 182
DB 115 IYANKLKQITIKIDGGEIAPNDPTNGRALIILQTAGLIKVDPAKQQLPTVSDITENKR 174
QY 183 KLVYKEVDTVAARAIDVDLAVVNNNYAGQVGLTASENGVEFVE---DKDSPYNIIVAR 239
DB 175 QLKITELDATQTAARALQVDVASVNSGAVDAGYTPDKDAIFLEPVNEKAKPYNIIVAR 234
QY 240 ADNKSRAIDQFVKAYOTDEVEAEAKKQFQGVIKGW 276
DB 235 EBDQENKLYQKVVEEYQOEETKVIATETSKGANVPAM 271

Search completed: June 16, 2004, 11:06:40
Job time : 63 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 16, 2004, 11:00:51 ; Search time 18 Seconds
(without alignments)
798.409 Million cell updates/sec

Title: US-10-018-672-2
Perfect score: 1383
Sequence: 1 MNFKINGICALLASGIALAG.....TDEVEAKKQKQGVINGW 276

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues
Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1012.5	73.2	277	1 PLPA_PASHA	Q08668 pasteurella
2	1006.5	72.8	273	1 METQ_HAEIN	P31728 haemophilus
3	888	64.2	276	1 METQ_PASWU	Q9CK95 pasteurella
4	793.5	57.4	271	1 METQ_SALUT	Q8Z932 salmonella
5	792.5	57.3	271	1 METQ_SALUT	Q8Z931 salmonella
6	791.5	57.2	263	1 PLPC_PASHA	Q08670 pasteurella
7	781.5	57.2	271	1 METQ_ECO57	Q8X8V9 escherichia
8	782.5	56.6	271	1 METQ_ECO57	P28635 escherichia
9	773	55.9	276	1 PLPB_PASHA	Q08669 pasteurella
10	757.5	54.8	271	1 METQ_YERPE	Q8Z440 yersinia pe
11	751.5	54.3	269	1 NLEA_ECOLI	Q9KJ17 vibrio chol
12	691.5	50.0	272	1 NLEA_ECOLI	P04846 escherichia
13	686.5	49.6	272	1 NLEA_ECOLI	Q8X550 escherichia
14	416.5	30.1	268	1 TP32_TREBA	Q07850 treponema p
15	286	21.4	263	1 YHCD_BACSU	P34594 bacillus su
16	112.5	8.1	267	1 RBP2_PLAVB	Q00799 plasmodium
17	106	7.7	1592	1 GTF2_STRDO	P27470 streptococc
18	104.5	7.6	991	1 SC44_RICAU	Q9A164 rickettsia
19	103.5	7.5	543	1 CH60_MYCCE	P47632 mycoplasma
20	103	7.4	1597	1 GTR1_STRDO	Q19001 streptococc
21	99.5	7.2	465	1 APRA_CIOAB	Q97K30 clostridium
22	99	7.2	814	1 SLA1_BACAA	P49051 bacillus an
23	99	7.2	928	1 HXA2_HAEIN	P45354 haemophilus
24	98.5	7.1	459	1 TRWE_CIOCE	Q89882 clostridium
25	98	7.1	469	1 Y889_METUA	Q45452 bacillus su
26	97.5	7.0	306	1 OPEC_BACSU	P47653 mycoplasma
27	97.5	7.0	1164	1 Y414_MYCCE	Q8E894 oesombacill
28	97	7.0	1366	1 Y414_MYCCE	Q12680 saccharomyc
29	96.5	7.0	339	1 PUR3_OCEIH	P52183 homo sapien
30	96.5	7.0	2144	1 GLU1_YEAST	P52312 campylobact
31	96	6.9	382	1 VATC_HUMAN	Q05038 ratobacter
32	96	6.9	428	1 AROA_CAMUE	
33	96	6.9	525	1 SPI_BAFPA	

34	96	6.9	546	1 CH60_AZOV1	Q8GBD4 azobacter
35	96	6.9	1729	1 RRP5_YEAST	Q05022 saccharomyc
36	95.5	6.9	431	1 AROA_AQUAE	Q67494 aquifex aeo
37	95.5	6.9	684	1 FLID_HELPY	P67866 helicobacte
38	95.5	6.9	1013	1 SC44_RICRH	Q9A181 rickettsia
39	95.5	6.9	1409	1 HAP1_HAEIN	P44596 haemophilus
40	95	6.9	382	1 VATC_MOUSE	Q92193 mus musculu
41	95	6.9	471	1 UDPG_ASTWE	Q91K97 astrogalus
42	95	6.9	513	1 LEU1_LACTA	Q02141 lactococcus
43	95	6.9	541	1 CH62_STRCO	Q9KXU5 streptococc
44	95	6.9	559	1 ILVD_PROMP	Q7V1C1 prochloroco
45	94.5	6.8	919	1 ATC1_RAT	Q64566 rattus norv

ALIGNMENTS

RESULT 1

ID	PLPA_PASHA	STANDARD	PRT	277 AA.
AC	Q08668; Q07363;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Outer membrane lipoprotein 1 precursor (PLP1).			
GN	PLPA.			
OS	Pasteurella haemolytica.			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;			
OC	Pasteurellaceae; Mannheimia.			
OX	NCBI_TaxID=75985;			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Serotype A1;			
RX	MEDLINE=9332810; PubMed=8335249;			
RA	Murphy G.V., Whitworth L.C.;			
RT	"Analysis of tandem, multiple genes encoding 30-kDa membrane proteins			
RT	in Pasteurella haemolytica A1."			
RL	Gene 129:107-111 (1993).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Serotype A1;			
RX	MEDLINE=94011378; PubMed=8405866;			
RA	Cooney B.J., Lo R.Y.C.;			
RT	"Three contiguous lipoprotein genes in Pasteurella haemolytica A1			
RT	type b."			
RL	Infect. Immun. 61:4682-4688 (1993).			
RT	-1- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid			
CC	anchor (probable).			
CC	-1- SIMILARITY: Belongs to the nlpA lipoprotein family.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; L11037; AAA25538.1; -			
DR	EMBL; M91072; AAA25541.1; -			
DR	EMBL; L16627; AAA25546.1; -			
DR	PIR; JN0751; JN0751.			
DR	InterPro; IPR004872; Lipoprotein_9.			
DR	InterPro; IPR000437; Prok_lipoprot_5.			
DR	InterPro; IPR004478; YaeC.			
DR	Pfam; PF03180; Lipoprotein_9; 1.			
DR	TIGRFAMs; TIGR00363; TIGR00363; 1.			
DR	TIGRFAMs; TIGR00363; TIGR00363; 1.			
DR	PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.			
KW	Outer membrane; Lipoprotein; Signal; Multigene family; Palmitate.			
FT	SIGNAL	1	19	POTENTIAL.
FT	CHAIN	20	277	OUTER MEMBRANE LIPOPROTEIN 1.
FT	LIPID	20	20	N-palmitoyl cysteine (Probable)

GN METO OR PLPA OR PM1730.
 OS Pasteurella multocida.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Pasteurella.
 OX NCBI_TaxID=747;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PM70;
 RX MEDLINE=21145866; PubMed=11248100;
 RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whitam T.S., Kapur V.;
 RT "Complete genomic sequence of Pasteurella multocida PM70.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
 CC -1- FUNCTION: This protein is a component of a D-methionine
 permease, a binding protein-dependent, ATP-driven transport
 system (By similarity).
 CC -1- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
 anchor (Probable).
 CC -1- SIMILARITY: Belongs to the alpha lipoprotein family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 DR EMBL; AEO06210; AK03814.1; -;
 DR InterPro; IPR004872; Lipoprotein_9.
 DR InterPro; IPR000437; Prok_lipoprot_S.
 DR InterPro; IPR004478; Yaec.
 DR Pfam; PF03180; Lipoprotein_9; 1.
 DR TIGRPFAMs; TIGR00363; TIGR00363; 1.
 DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.
 DR Transport; Amino-acid transport; Outer membrane; Lipoprotein; Signal;
 KM Complete proteome; Palmitate.
 FT SIGNAL 1
 FT CHAIN 1 20 PROBABLE D-METHIONINE-BINDING LIPOPROTEIN
 FT LIPID 21 21 N-palmitoyl cysteine (Probable).
 FT LIPID 21 21 S-diacylglycerol cysteine (Probable).
 SQ SEQUENCE 276 AA; 30232 MW; 7AFD862A687D624 CRC64;

 Query Match 64.2%; Score 888; DB 1; Length 276;
 Best Local Similarity 60.2%; Pred. No. 1.3e-52;
 Matches 168; Conservative 55; Mismatches 50; Indels 6; Gaps 3;
 OY 1 MNFGINGICALASGIALAGCSNOSNEBPAIS---KTAAGITIKYVAGPQAAVAEVAQ 57
 DB 1 MTLTFLGELATLIVAVALLAGC--KODKPAALAAAPSPAPARKLTGVWMTGAAGQTEVAAK 58
 OY 58 VAKERTNLTVLVEENDYAMPNSAVSKGLDANNMOKHPYLEKDSQEKGLNLTIVGNTF 117
 DB 59 IAKEXKINDVLTVEFTETQFNDAITKGLDANNAFQHPWMDKVEGRGY-KLAIYGNTF 117
 OY 118 VYPLAGVSTKTKTINELDQATIVAPNPSULARALLILEQGLIKKQNTNLFSTLIDI 177
 DB 118 VFPPLAASKTKIKNVSELDGATVAAPNPNISGLRALLILEQGLIKKQDPSNLFSTSIDV 177
 OY 178 VENPKVLVIKEVDIVSAAARLIDVDLAVANNVAGQVGLTASENGVFVEDDSFYVNIIV 237
 DB 178 IENPKNLDIKVEVSLPRMLDDVDFAINNNVAVQGLTKEKGIIVEDDSFYVNIIV 237
 OY 238 ARAADKSKALQDPYKAVQTDVEVAEAKKQFKQGIYKGM 276
 DB 238 SREDNKDNEALKDFKVAQTBEVYQEALEKHPQGGVYVGM 276

DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE D-methionine-binding lipoprotein metc precursor.
 GN METO OR STY0272 OR T0248.
 OS Salmonella typhi.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=601;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CT18;
 RX MEDLINE=21534947; PubMed=11677608;
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahinia M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 RA Feltwell T., Hamlin N., Hague A., Hen T.T., Holroyd S., Jorgels K.,
 RA Krogh A., Larsen T.S., Leather S., Moule S., O'Garra P., Parry C.,
 RA Quail M.A., Rutherford K., Simmonds M., Skellon J., Stevens K.,
 RA Whitehead S., Barrett B.G.;
 RT "Complete genome sequence of a multiple drug resistant Salmonella
 enterica serovar Typhi CT18.";
 RL Nature 413:848-852(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TY2 / ATCC 700931;
 RX MEDLINE=22531367; PubMed=12644504;
 RA Deng W., Lion S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
 RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
 RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
 and CT18.";
 RL J. Bacteriol. 185:2330-2337(2003).
 CC -1- FUNCTION: This protein is a component of a D-methionine
 permease, a binding protein-dependent, ATP-driven transport
 system (By similarity).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 (Probable).
 CC -1- MISCELLANEOUS: The metNIO system is also to be able to transport
 the toxic methionine analog alpha-methyl-methionine (By
 similarity).
 CC -1- SIMILARITY: Belongs to the alpha lipoprotein family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 DR EMBL; AL627266; CAD08705.1; -;
 DR EMBL; AEO16834; AAC67977.1; -;
 DR InterPro; IPR004872; Lipoprotein_9.
 DR InterPro; IPR000437; Prok_lipoprot_S.
 DR InterPro; IPR004478; Yaec.
 DR Pfam; PF03180; Lipoprotein_9; 1.
 DR TIGRPFAMs; TIGR00363; TIGR00363; 1.
 DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.
 DR Transport; Amino-acid transport; Membrane; Lipoprotein; Signal;
 KM Complete proteome; Palmitate.
 FT SIGNAL 1
 FT CHAIN 1 22 POTENTIAL.
 FT LIPID 23 23 D-METHIONINE-BINDING LIPOPROTEIN METO.
 FT LIPID 23 23 N-palmitoyl cysteine (Potential).
 FT LIPID 23 23 S-diacylglycerol cysteine (Potential).
 SQ SEQUENCE 271 AA; 29465 MW; 49B62C4CF96D9613 CRC64;

 Query Match 57.4%; Score 793.5; DB 1; Length 271;
 Best Local Similarity 56.5%; Pred. No. 2.6e-46;
 Matches 156; Conservative 44; Mismatches 65; Indels 11; Gaps 3;
 OY 3 FGKINGICALASGIALAGCSNOSNEBPAISKTAAGITIKYVAGPQAAVAEVAQVAKEX 62
 DB 5 FKTPAAGVGLISGLAGCGDDEKDP-----NHIKGVIVGAEEQVAAEVAQVAKEX 56

```

QY 63 YNLVLEVEFNDYAMPNSAVSGEILDANAMQHKPELEKDSQEGKLNVLVGTTFYVPLA 122
DB 57 YGLDVELVTNDYVLPNAAASKGDIIDANAFQHKPYLDQIKRGY-KLVSVGKTFYVPLA 115
QY 123 GYSTKIKTLNELKOGATIAVNDPSNLARALLILEKGLIKDKNTNLTSTTLDIYENK 182
DB 116 GYSKIKSIDELKDSQVAVPNDPTNLGRSLLLQKVGILLKDKDVGSLPTSLDIYENK 175
QY 183 KLVIKEDVTSVAARAIDD--VDLAVVNNNYAGVGLTASENGVEEDKDSFYVNIIVARA 240
DB 176 NLKIVLEAPQLPRSLDDAQIALAVINTTYASQIGLTPAKDGI FVEDKDSFYVNIIVTRE 235
QY 241 DNKSKAIODFVAKYQTDVEVAEAKKQFQVIGKM 276
DB 236 DNKDAENVKKEFVQAYQSDVEVAEANKVFNGAVKGM 271

RESULT 5
METO_SALTY
ID METO_SALTY STANDARD; PRT; 271 AA.
AC 082RNL;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DB D-methionine-binding lipoprotein meto precursor.
GN METO OR STM0245.
OS Salmoneella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmoneella.
OX NCBI_TaxID=602;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=L72 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Rader S., Sun H., Florea L., Miller W., Stuckert T., Nian M.,
RA Watson R., Wilson R.K.;
RT "Complete genome sequence of Salmoneella enterica serovar Typhimurium
RT L72."
RL Nature 413:852-856 (2001).
CC -!- FUNCTION: This protein is a component of a D-methionine
CC permease, a binding protein-dependent, ATP-driven transport
CC system (by similarity).
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC (Probable).
CC -!- MISCELLANEOUS: The metNIQ system is also able to transport
CC the toxic methionine analog alpha-methyl-methionine (by
CC similarity).
CC -!- SIMILARITY: Belongs to the alpha lipoprotein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL; AF008706; AAL19208.1; -
DB StGene; SG7777; metQ.
DB InterPro; IPR004872; Lipoprotein_9.
DB InterPro; IPR000437; Prok_Lipoprot_S.
DB InterPro; IPR004478; YaeC.
DB Pfam; PF01180; Lipoprotein_9; 1.
DB TIGRfams; TIGR00363; TIGR00363; 1.
DB PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
DB Transport; Amino-acid transport; Membrane; Lipoprotein; Signal;
DB Complete proteome; Palmitate.
DB SIGNAL 1 22
DB CHAIN 23 271 D-METHIONINE-BINDING LIPOPROTEIN METO.

```

```

FT LIPID 23 23 N-palmitoyl cysteine (Potential).
FT LIPID 23 23 S-diacylglycerol cysteine (Potential).
SQ SEQUENCE 271 AA; 29435 MW; 3F6D41E14FBD707 CRC64;

Query Match.
Best local similarity 56.5%; Score 792.5; DB 1; Length 271;
Matches 156; Conservative 43; Mismatches 66; Indels 11; Gaps 3;

QY 3 FGKINGICALASGIALAGCSNOSNEPAISKTAQTIKGVNAGSEQA VAEVAGVAREK 62
DB 5 FKTPAAVAGLIGSLALAGCGODEKDP-----NHIKGVIVGAEOQVAEVAQVAREK 56
QY 63 YNLVLEVEFNDYAMPNSAVSGEILDANAMQHKPELEKDSQEGKLNVLVGTTFYVPLA 122
DB 57 YGLDVELVTNDYVLPNAAASKGDIIDANAFQHKPYLDQIKRGY-KLVSVGKTFYVPLA 115
QY 123 GYSTKIKTLNELKOGATIAVNDPSNLARALLILEKGLIKDKNTNLTSTTLDIYENK 182
DB 116 GYSKIKSIDELKDSQVAVPNDPTNLGRSLLLQKVGILLKDKDVGSLPTSLDIYENK 175
QY 183 KLVIKEDVTSVAARAIDD--VDLAVVNNNYAGVGLTASENGVEEDKDSFYVNIIVARA 240
DB 176 NLKIVLEAPQLPRSLDDAQIALAVINTTYASQIGLTPAKDGI FVEDKDSFYVNIIVTRE 235
QY 241 DNKSKAIODFVAKYQTDVEVAEAKKQFQVIGKM 276
DB 236 DNKDAENVKKEFVQAYQSDVEVAEANKVFNGAVKGM 271

RESULT 6
PLPC_PASHA
ID PLPC_PASHA STANDARD; PRT; 263 AA.
AC 006870; Q07365;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DB Outer membrane lipoprotein 3 precursor (Plp3).
GN PlpC.
OS Pasteurella haemolytica.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Mannheimia.
OX NCBI_TaxID=75985;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Serotype A1;
RX MEDLINE=93328110; PubMed=8335249;
RA Murphy G.L., Whitworth L.C.;
RT "Analysis of tandem, multiple genes encoding 30-kDa membrane proteins
RT in Pasteurella haemolytica A1."
RL Gene 129:107-111 (1993).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=Serotype A1;
RX MEDLINE=84011378; PubMed=8406866;
RA Cooney B.J., Lo R.Y.C.;
RT "Three contiguous lipoprotein genes in Pasteurella haemolytica A1
RT which are homologous to a lipoprotein gene in Haemophilus influenzae
RT type b."
RL Infect. Immun. 61:4682-4688 (1993).
CC -!- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
CC anchor (Probable).
CC -----
CC -!- SIMILARITY: Belongs to the alpha lipoprotein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL; L11037; AAA25540.1; -
DB EMBL; L16627; AAA25548.1; ALT_INIT.

```


DR	EMBL; AE000129; AAC73308.1; -.
DR	EMBL; D93523; BAA78784.1; -.
DR	EMBL; U07214; AAB06825.1; -.
DR	EMBL; L04474; AAA24507.1; ALT_FRAME.
DR	PIR; E64744; E64744.
DR	SWISS-2DPAGE; P28635; COLI.
DR	EcoGene; EGJ1504; metC.
DR	InterPro; IPR004872; Lipoprotein_9.
DR	InterPro; IPR000437; Prok_lipoprot_s.
DR	InterPro; IPR004478; YaeC.
DR	pfam; PF03180; Lipoprotein_9; 1.
DR	TIGRFAMs; TIGR00363; tIGR00363.1.
DR	PROSITE; PS00013; PROKAR_LIPOPROTEIN_1.
KW	Transport; Amino-acid transport; Membrane; Lipoprotein; Signal;
KW	Complete proteome; Palmitate.
FT	SIGNAL 1 22
FT	CHAIN 23 271
FT	LIPID 23 23
FT	LIPID 23 23
FT	CONFLICT 200 200
SO	SEQUENCE 271 AA; 29431 MW; B50CB6FB95CD2EB7.CRC64;
Query Match	56.6%; Score 782.5; DB 1; Length 271;
Best Local Similarity	55.1%; Pred. No. 1.4e-45;
Matches 152; Conservative	46; Mismatches 67; Indels 11; Gaps 3
Dy	3 FGGKINCITGLAGSGLACGNSQSNPEPAISKTAAQTIKGVGMGPDAVAEYVAGOVAEXK 62
Dy	5 FKTFPAAVGALISGLVLVCQGDEKDP-----NHKIVGVGAHQVAVAEVAKYAKDK 56
Oy	63 YNLTVELVEFNDYAMPNSAVSKGELDANAMOHKEYLEKDSQEGKLNNLVIQNTVPYPLA 122
Dy	57 YGDLVELVFENDYVLPNPAISKDIDANAFAHKKPYLDQQAKDRGY-KLVAVGNTEFYYPYA 115
Oy	123 GYSTTKIKTLNELKDGATTA VNDPSNLAPRLILLEKGKIKLKMDNNLTSTLLDIYENPK 182
Dy	116 GYSKIKKLSIDELODSDQVAVPNDDPTNGRSLLILKQKGLKLDGVGLPLTVDVVENPK 175
Oy	183 KLVIEKVDPTSVAARAIDD--VDLA VVNNTNYAGQVGLTASENGYVEDKDSPVYNIIIVARA 240
Dy	176 NIKLVIELAEPLQLRPBRLDAQIALVINTVTYASQIGLTPADGIFVEDKESPVYNILVIRE 235
Oy	241 DNKDSKAIDPFYKAYQTDVEVAEKQKQFKQGVKGM 276
Dy	236 DNKDAAENVKKEFQAQYOSDEVYEAAANKYFNCGAAVKGM 271
RESULT 9	
ID	PIPB_PASHA STANDARD; PTR; 276 AA.
AC	008859; O07364;
DT	01-NOV-1995 (Rel. 32, Created)
DT	01-NOV-1995 (Rel. 32, Last sequence update)
DT	10-OCT-2003 (Rel. 42, Last annotation update)
D5	Outer membrane lipoprotein 2 precursor (Plp2).
GN	PIPB.
OS	Pasteurella haemolytica.
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
CC	Pasteurellaceae; Mannheimia.
CX	NCBI_TaxId=75985;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=serotype A1;
RC	MEDLINE=93328110; PubMed=8335249;
RA	Murphy G.L., Whitworth L.C.;
RT	"Analysis of tandem, multiple genes encoding 30-kDa membrane proteins
RL	in Pasteurella haemolytica A1.";
LN	Gene 129:107-111(1993).
RE	[2]
RC	SEQUENCE FROM N.A.
RC	STRAIN=serotype A1;
RC	MEDLINE=94011378; PubMed=8406866;
RA	Cooney B.U., Lo R.Y.C.;

RT "Three contiguous lipoprotein genes in *Pasteurella haemolytica* A1
 RT type b.",
 RT which are homologous to a lipoprotein gene in *Haemophilus influenzae*
 RL Infect. Immun. 61:4692-4698(1993).
 CC - SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
 CC anchor (Probable).
 CC - SIMILARITY: Belongs to the alpha lipoprotein family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L11037; AAA25539.1; -
 DR EMBL: L16627; AAA25547.1; -
 DR PIR: JN0752; JN0752.
 DR InterPro: IPR004872; Lipoprotein 9.
 DR InterPro: IPR000437; Prok lipoprot_5.
 DR InterPro: IPR004478; YaeC
 DR Pfam: PF03180; Lipoprotein 9; 1.
 DR TIGRfam: TIGR00363; TIGR00363; 1.
 DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 DR Outer membrane; Lipoprotein; signal; Multigene family; Palmitate.
 KW SIGNAL. 1
 FT CHAIN 20 276 OUTER MEMBRANE LIPOPROTEIN 2.
 FT LIPID 20 20 N-palmitoyl cysteine (Probable).
 FT LIPID 20 20 S-diacylglycerol cysteine (Probable).
 FT CONFLICT 268 270 FNG -> LVH (IN REF. 2).
 FT SEQUENCE 276 AA; 30157 MW; DF8P4434ADAD950 CRC64;
 SQ
 Query Match 55.9%; Score 773; DB 1; Length 276;
 Best Local Similarity 54.0%; Pred. No. 6,3e-45;
 Matches 150; Conservative 54; Mismatches 70; Indels 4; Gaps 4;
 QY 1 MNFGKINGICALASGIALAGCSNQ-SNEPAA:SKTAQT-IRVGVWAGEQVAEVAQGV 58
 DB 1 MNFGKINGICALASGIALAGCSNQ-SNEPAA:SKTAQT-IRVGVWAGEQVAEVAQGV 58
 QY 1 MNFGKINGICALASGIALAGCSNQ-SNEPAA:SKTAQT-IRVGVWAGEQVAEVAQGV 58
 DB 1 MNFGKINGICALASGIALAGCSNQ-SNEPAA:SKTAQT-IRVGVWAGEQVAEVAQGV 58
 QY 59 AAEKXVLTVEIVFENDYAMPNSAVSGEIDANAMQKPYLEDSOEKGLNNTVIYGNFTV 118
 DB 60 AAEKXVLTVEIVFENDYAMPNSAVSGEIDANAMQKPYLEDSOEKGLNNTVIYGNFTV 118
 QY 119 YPLAGYSTIKITLNEKDGATIAVPDPSNLARALILEKQGLIKLKDNTLFTLLDIV 178
 DB 119 YPLAGYSTIKITLNEKDGATIAVPDPSNLARALILEKQGLIKLKDNTLFTLLDIV 178
 QY 119 WPIAAVSKIKIKNSKELKDGATVAPNNASNTARALLLGAHGLIKLKDKPFAIENDII 178
 DB 119 WPIAAVSKIKIKNSKELKDGATVAPNNASNTARALLLGAHGLIKLKDKPFAIENDII 178
 QY 179 ENPKKLVIEVNTSVAAARAIIDVDLAVNNNTNAGVGLTASENGVFVEKDSPPYNIYA 238
 DB 179 ENPKKLVIEVNTSVAAARAIIDVDLAVNNNTNAGVGLTASENGVFVEKDSPPYNIYA 238
 QY 239 RADNKSKAIDPFVAKYQTFVEAEAKQFQKGVIKGM 276
 DB 239 RADNKSKAIDPFVAKYQTFVEAEAKQFQKGVIKGM 276
 QY 239 REDNNDPRLQITFVASFOTEIVFGEALKLFNGVYKGM 276
 DB 239 REDNNDPRLQITFVASFOTEIVFGEALKLFNGVYKGM 276
 RESULT 10
 METQ_YERPE STANDARD; PRT; 271 AA.
 ID METQ_YERPE
 AC 082H40;
 DT 28-FEB-2003 (Rel. 41; Created)
 DT 28-FEB-2003 (Rel. 41; Last sequence update)
 DT 10-OCT-2003 (Rel. 42; Last annotation update)
 DE D-methionine-binding lipoprotein metQ precursor.
 GN METQ OR YP01071 OR Y3106.
 OS Yersinia pestis.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Yersinia.
 OC NCBI_Taxid=632;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-CO-92 / Bivox Orientalis;
 RX MEDLINE=21470413; PubMed=11586360;
 RA Parthill U., Wren B.W., Thomson N.R., Tibball R.W., Holden M.T.G.,
 RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
 RA Felwell T., Hamlin N., Holroyd S., Jagels K., Karylyuk A.V.,
 RA Leather S., Meule S., Oyston P.C.F., Quail M.A., Rutherford K.,
 RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrett B.G.,
 RT Genome sequence of *Yersinia pestis*, the causative agent of plague.";
 RL Nature 413:523-527(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-KIMS / Bivox Mediaevalis;
 RX MEDLINE=22137863; PubMed=12142430;
 RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
 RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
 RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
 RA Straley S.C., McDonough K.A., Nilles M.L., Watson J.S., Blattner F.R.,
 RA Perry R.D.;
 RT "Genome sequence of *Yersinia pestis* KIM.";
 RL J. Bacteriol. 184:4601-4611(2002).
 CC - FUNCTION: This protein is a component of a D-methionine
 CC permease, a binding protein-dependent, ATP-driven transport
 CC system (By similarity).
 CC - SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 CC (Probable).
 CC - MISCELLANEOUS: The metIQ system is also to be able to transport
 CC the toxic methionine analog alpha-methyl-methionine (By
 CC similarity).
 CC - SIMILARITY: Belongs to the alpha lipoprotein family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AJ141416; CAC99914.1; -
 DR EMBL: AE013911; AA06656.1; -
 DR PIR: AG0131; AG0131.
 DR InterPro: IPR004872; Lipoprotein 9.
 DR InterPro: IPR000437; Prok lipoprot_5.
 DR InterPro: IPR004478; YaeC
 DR Pfam: PF03180; Lipoprotein 9; 1.
 DR TIGRfam: TIGR00363; TIGR00363; 1.
 DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 DR Transport; Amino-acid transport; Membrane; Lipoprotein; signal;
 KW Complete proteome; Palmitate.
 KW SIGNAL. 1
 FT CHAIN 23 271 D-METHIONINE-BINDING LIPOPROTEIN METQ.
 FT LIPID 23 23 N-palmitoyl cysteine (Potential).
 FT LIPID 23 23 S-diacylglycerol cysteine (Potential).
 FT SEQUENCE 271 AA; 29576 MW; 5214CEB0517FC8 CRC64;
 SQ
 Query Match 54.8%; Score 757.5; DB 1; Length 271;
 Best Local Similarity 53.2%; Pred. No. 6,7e-44;
 Matches 148; Conservative 46; Mismatches 71; Indels 11; Gaps 3;
 QY 1 MNFGKINGICALASGIALAGCSNQ-SNEPAA:SKTAQT-IRVGVWAGEQVAEVAQGV 60
 DB 3 LKFSIAIAISLIGTLVLVGGPTKAP-----NHIVGVIVGEQVAAEVAQGVAK 54
 QY 61 EKXVLTVEIVFENDYAMPNSAVSGEIDANAMQKPYLEDSOEKGLNNTVIYGNFTV 120
 DB 55 EKXVLTVEIVFENDYAMPNSAVSGEIDANAMQKPYLEDSOEKGLNNTVIYGNFTV 113
 QY 121 LAGYSTIKITLNEKDGATIAVPDPSNLARALILEKQGLIKLKDNTLFTLLDIVEN 180
 DB 114 LAGYSTIKITLNEKDGATIAVPDPSNLARALILEKQGLIKLKDNTLFTLLDIVEN 173

106 JUN 22 11:03:07 2002

QY 181 PKTLVKEVDTSVAARAID--VDLAVNNNNAAGVGLTASENGVEPEDSDPYNNIYA 238
 DB 174 PKNLKVELEAPLPSPDLDQIALAIINTTAAQGLTPADGLFVEDKESPYNLLVA 233
 QY 239 RADNDSKAIQDFKAYQTDVEAEAKQFKQGVIKGM 276
 DB 234 REDNKAENKRFVQAYQSDVEYDAANKAFNGAVKGM 271

RESULT 11
 ID METO_VIBCH STANDARD; PRT; 269 AA.
 AC 09KTU7;
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Probable D-methionine-binding lipoprotein metQ precursor.
 GN METQ OR VC0305.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 CC Vibrionaceae; Vibrrio.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EI TOR N16961 / Serotype O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Bisen J.A., Nelson M.C., Clayton R.A., Gwin M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Esmolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L., Ueberack T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.D., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 cholerae";
 RL Nature 406:477-483(2000).
 CC -1- FUNCTION: This protein is a component of a D-methionine
 permease, a binding protein-dependent, ATP-driven transport
 system (By similarity).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 (Probable).
 CC -1- SIMILARITY: Belongs to the nlpa lipoprotein family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL; AE004174; AAF94067.1; ALT_INIT.
 DR TIGR; VC0905; -
 DR InterPro; IPR004872; Lipoprotein_9.
 DR InterPro; IPR000437; Prok_lipoprot_S.
 DR InterPro; IPR004478; YaeC
 DR Pfam; PF03180; Lipoprotein_9; 1.
 DR TIGRFAMs; TIGR00363; TIGR00363; 1.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 DR TRANSPORT; Amino-acid transport; Membrane; Lipoprotein; Signal;
 KW Complete proteome; Palmitate.
 FT SIGNAL 1 22
 FT CHAIN 23 269
 FT LIPID 23 23 N-palmitoyl cysteine (potential).
 FT LIPID 23 23 S-diacylglycerol cysteine (potential).
 SQ SEQUENCE 269 AA; 29067 MW; 39E2570BE3F184D6 CRC64;

Query Match 54.3%; Score 751.5; DB 1; Length 269;
 Best Local Similarity 54.1%; Pred. No. 1,7e-43;
 Matches 145; Conservative 51; Mismatches 63; Indels 9; Gaps 2;

QY 9 ICALASGIALAGSGNSNEPAISKTAQTIKVWAGPEQVAEVAAGVAKKINLTVE 68

DB 11 IAAASLTLLAGGEKAVD-----NNKVKIGWAGAAQVAEVAAKAKEKYLIDVE 62
 QY 69 LVFENDYAMNSAVSKKELDANAMQHPRYLEKSOEGKLNNTYVNGTFPYPLAGSTKI 126
 DB 63 LVTFPTDYVTPAALDDSIDANAFQHKPYLDKCIABRGY-KLAIVGTFVYPIAGISKQI 121
 QY 129 KTLNELKQATIAVNPDPNSMLARALILKEQGIKDKONTNLTSTLDIVENPKLVYKE 168
 DB 122 KSVLELDQGARIVNPDPNTLGRSLILLLQGGILKLRBDVGLLATVADIVENPKLILE 161
 QY 189 VDISVAARAIDVDLAVNNNNAAGVGLTASENGVEPEDSDPYNNIYVARANKSKAI 248
 DB 182 LDAQQLPRSLDDVALSTINTTYASSINLTPEKQGI FVENKESPYNLLVAREANVAENV 241
 QY 249 QDFKAYQTDVEAEAKQFKQGVIKGM 276
 DB 242 QNFKKAQTDVEAKAAAEITFGGAVKGM 269

RESULT 12
 ID NLPA_ECOLI STANDARD; PRT; 272 AA.
 AC P04846;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Lipoprotein-28 precursor.
 GN NLPA OR B3661.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KL2 / MG1655;
 RX MEDLINE=93315143; PubMed=7686882;
 RA Burtand V.D., Plunkett G. III, Daniels D.L., Blattner F.R.;
 RA "DNA sequence and analysis of 136 kilobases of the Escherichia coli
 genome: organizational symmetry around the origin of replication";
 RT Genomics 16:551-561(1993).
 CC -1- SUBCELLULAR LOCATION: Attached to the inner membrane by a lipid
 anchor.
 CC -1- SIMILARITY: Belongs to the nlpa lipoprotein family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL; M12163; AAA24080.1; -
 DR EMBL; L10328; AAA62013.1; -
 DR EMBL; AE000443; AAC76684.1; -
 DR PIR; A26286; LPEC28.
 DR SWISS-2DPAGE; P04846; COLI.
 DR Ecocore; EG10657; nlpa.
 DR InterPro; IPR004872; Lipoprotein_9.
 DR InterPro; IPR000437; Prok_lipoprot_S.
 DR InterPro; IPR004478; YaeC
 DR Pfam; PF03180; Lipoprotein_9; 1.
 DR TIGRFAMs; TIGR00363; TIGR00363; 1.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 KW Inner membrane; Lipoprotein; Signal; Complete proteome; Palmitate.

FT SIGNAL 1 23
 FT CHAIN 24 272 LIPOPROTEIN-28.
 FT LIPID 24 24 N-palmitoyl cysteine.
 FT LIPID 24 24 S-diacylglycerol cysteine.
 SQ SEQUENCE 272 AA; E3C249E753AB1B33 CRC64;

Query Match 50.0%; Score 691.5; DB 1; Length 272;
 Best Local Similarity 50.6%; Pred. No. 1,8e-39;
 Matches 137; Conservative 53; Mismatches 70; Indels 11; Gaps 3;

QY 8 GICALASGIALAGCSNOSNEPAISKTAAQTIVGWAGPEQAAVEAGVAKKNTLV 67
 DB 11 GAALLAGILLAGCQSSD-----AKHIVGVINGAEQDAVEAKKVAKEKYGLDV 62
 QY 68 ELVEENDYAMPNSAVSKGELDANAMQHKPYLEKDSQELNNLVIVGNTFYVPLAGVSTK 127
 DB 63 ELVGFSGSLPNDATNGHLDANVFQHRPFLQDNQAHGY-KLVAAGNTFVFPWAGYSKK 121
 QY 128 IKTLNELKDGATIAVPNDPSNLARALLILEKQGLIKLKDNTNLFSTLIDIVENPKVLV 187
 DB 122 IKTVAQIKEGATVAILPNDPTNLGRALLLLQKEKILITLKEGKGLPTALDITDNPRLQIM 181
 QY 188 EVDTSVAAARAID--VDLAVNNNYAGQVGLTASENGVFVEKDSPPYNIIVARADNDS 245
 DB 182 ELEGAQLPRLVDDPKVDVALISTTYIQGTGLSPVHDSVFIEDKNSPYNIIIVAREDNKA 241
 QY 246 KAIDPFVAVOTDEVEAEAKKQFQDVIGKW 276
 DB 242 ENVKEFLOSYSQSPVAKAAETIFNGAVPGW 272

RESULT 13
 NIPA_ECO57
 ID NIPA_ECO57 STANDARD; PRT; 272 AA.
 AC 08XC50:
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Lipoprotein-28 precursor.
 GN NIPA OR Z5147 OR ECS4595.
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC Enterobacteriaceae; Escherichia.
 NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 Rose D.U., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller U.,
 Grobbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamowski K.,
 Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 Welch R.A., Blattner F.R.;
 RA Genome sequence of enterohemorrhagic Escherichia coli O157:H7.;
 RT Nature 409:529-533(2001).
 RL [2]

SEQUENCE FROM N.A.
 RP STRAIN=O157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; PubMed=1158796;
 RA Hayashi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,
 Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 Rida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
 Kihara S., Shida T., Hattori M., Shingawa H.,
 RT "Complete genome sequence of enterohemorrhagic Escherichia coli
 O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:11-22(2001).
 CC -1- SUBCELLULAR LOCATION: Attached to the inner membrane by a lipid
 anchor (By similarity).
 CC -1- SIMILARITY: Belongs to the nipa lipoprotein family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

DR EMBL; AE005597; AAG58856.1; -
 DR EMBL; AP002566; BAB38018.1; -
 DR PIR; C91203; C91203.
 DR PIR; D86049; D86049.
 DR InterPro; IPR004872; Lipoprotein_9.
 DR InterPro; IPR00437; Prok_Lipoprot_8.
 DR InterPro; IPR004478; Yaec.
 DR Pfam; PF03180; Lipoprotein_9; 1.
 DR TIGRFAMs; TIGR00363; TIGR00363; 1.
 DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.
 KM Inner membrane; Lipoprotein; Signal; Complete proteome; Palmitate.
 FT SIGNAL 1 23
 FT CHAIN 24 272 LIPOPROTEIN-28 (By similarity).
 FT LIPID 24 24 N-palmitoyl cysteine.
 FT LIPID 24 24 S-diacylglycerol cysteine.
 SQ SEQUENCE 272 AA; A52B4C8576B443F5 CRC64;

Query Match 49.6%; Score 686.5; DB 1; Length 272;
 Best Local Similarity 49.8%; Pred. No. 3.8e-39;
 Matches 135; Conservative 54; Mismatches 71; Indels 11; Gaps 3;

QY 8 GICALASGIALAGCSNOSNEPAISKTAAQTIVGWAGPEQAAVEAGVAKKNTLV 67
 DB 11 GAALLAGILLAGCQSSD-----EKHIVGVINGAEQDAVEAKKVAKEKYGLDV 62
 QY 68 ELVEENDYAMPNSAVSKGELDANAMQHKPYLEKDSQELNNLVIVGNTFYVPLAGVSTK 127
 DB 63 ELVGFSGSLPNDATNGHLDANVFQHRPFLQDNQAHGY-KLVAAGNTFVFPWAGYSKK 121
 QY 128 IKTLNELKDGATIAVPNDPSNLARALLILEKQGLIKLKDNTNLFSTLIDIVENPKVLV 187
 DB 122 IKTVAQIKEGATVAILPNDPTNLGRALLLLQKEKILITLKEGKGLPTALDITDNPRLQIM 181
 QY 188 EVDTSVAAARAID--VDLAVNNNYAGQVGLTASENGVFVEKDSPPYNIIVARADNDS 245
 DB 182 ELEGAQLPRLVDDPKVDVALISTTYIQGTGLSPVHDSVFIEDKNSPYNIIIVAREDNKA 241
 QY 246 KAIDPFVAVOTDEVEAEAKKQFQDVIGKW 276
 DB 242 ENVKEFLOSYSQSPVAKAAETIFNGAVPGW 272

RESULT 14
 TP32_TREPA
 ID TP32_TREPA STANDARD; PRT; 268 AA.
 AC 007950:
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Membrane lipoprotein Tpn32 precursor (29 kDa protein).
 GN TP32 OR TP0821.
 OS Treponema pallidum.
 OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
 NCBI_TaxID=160;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Porcella S.F., Radolf J.D., Norgard M.V.;
 RT "Treponema pallidum 29k protein is homologous to a lipoprotein
 present in Pasteurella hemolytica and in Hemophilus influenzae
 type D.";
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.

SEQUENCE FROM N.A.
 RP STRAIN=Nichols;
 RC Stamm L.V., Barnes N.Y.;
 RA "Treponema pallidum lipoprotein homologue";
 RT Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.

[3]
 RN SEQUENCE FROM N.A.
 RP STRAIN=Nichols;
 RX MEDLINE=98332770; PubMed=9665876;
 RA Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G.,
 RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,
 RA Sodergren E., Hardham J.M., McLeod M.P., Salberg S., Peterson J.,
 RA Kralak H., Richardson D., Howell J.K., Chidambaram M., Uterback T.,
 RA McDonald L., Atzich P., Bowman C., Cotton M.D., Fujii C., Garland S.,
 RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
 RA Venter J.C.;
 RT "Complete genome sequence of *Treponema pallidum*, the syphilis
 RT spirochete.";
 RL Science 281:375-388(1998).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 CC (Probable).
 CC -1- SIMILARITY: Belongs to the nlpa lipoprotein family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U97358; AAB63362.1; -;
 DR EMBL: U93844; AAB61267.1; -;
 DR EMBL: AF001253; AAC65789.1; -;
 DR FPIR: G71275; G71275.
 DR TIGR: TP0821; -;
 DR InterPro: IPR004872; Lipoprotein_9.
 DR InterPro: IPR000437; Prok_lipoprot_9.
 DR Pfam: PF03180; Lipoprotein_9; 1.
 DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 KW Membrane; Lipoprotein; Signal; Complete proteome; Palmitate.
 FT SIGNAL 1 23
 FT CHAIN 1 23 BY SIMILARITY.
 FT LIPID 24 24 MEMBRANE LIPOPROTEIN TP032.
 FT LIPID 24 24 N-palmitoyl cysteine (Potential).
 FT LIPID 24 24 S-diacylglycerol cysteine (Potential).
 SQ SEQUENCE 268 AA; 29081 MW; AB21F8A07A88691 CRG64;
 Query Match 30.1%; Score 416.5; DB 1; Length 268;
 Best Local Similarity 34.1%; Pred. No. 4,4e-21;
 Matches 92; Conservative 61; Mismatches 100; Indels 17; Gaps 4;
 QY 4 GKNIGICALASGIALAGCSNOSNEPAISKTAOTIKYGVNAGPQAAVEAGVAKEXY 63
 DB 12 GKL-----IALSGVAVACTGVDE-----IVGVGVSEPHARKLLEIAKEVYK-KQ 55
 QY 64 NLTVLEVENDYAMNSAVSKGELDANAMQHKPYLEKDSQEKGLNNLVIVGTFYPIAG 123
 DB 56 HIELRLVEFTNYVALNEAVMRGDIIMNFQHVPHMQFQOEHN-GDLVSGVGHVEPLAL 114
 QY 124 YSTKTKLTNELKDGATIAVNDPSNLARLTILEKQGLIKDKNTNLFSTLLDIYENPK 183
 DB 115 YSRTRHVSDFPAGAVIATIPDSNEARLRLLLEAAGFIRKAGSGGLFATVEEDVOQVKN 174
 QY 184 LVIKEDYTSVAAARAIDVDLAVNNNYAGQVGLTASENGVEVEDSPFYVNIIVARADNK 243
 DB 175 VVLQEVESALLPRVFDQVGVAVINGVYIMAGLSARBDGLAVEPDASAVAVLVYKRGNE 234
 QY 244 DSKATIDPFYKATQDEVEAEAKQKRDGVI 273
 DB 235 ADARVQAVIRALCGGRVRTYIKERYKGEV 264
 RESULT 15
 YHCU BACSU STANDARD; PRT; 263 AA.
 AC PS4534;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)

10-OCT-2003 (Rel. 42, Last annotation update)
 DE Hypothetical lipoprotein yhcJ precursor.
 GN YHCU OR BSU09110.
 OS Bacillus subtilis.
 CC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=168;
 RA Nobsack M.A., Tepstra P., Holsappel S., Venema G., Bron S.;
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 RP SEQUENCE FROM N.A.
 RA STRAIN=168 / JH642;
 RA Wendrich T.M., Marahiel M.A.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 RP SEQUENCE FROM N.A.
 RP STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertolo M.G., Bessieres P., Boloitin A., Borchert S.,
 RA Boursier R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Comerion I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Funa S., Gallizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
 RA Guisepi G., Guy B.J., Haga K., Haeisch J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hulle M.F., Itaya M., Jones L.,
 RA Jorja B., Karamata D., Kasehara Y., Klaeer-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kunano M.,
 RA Kurita K., Lapius A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Mostl D., Nakai S., Nobsack M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Odega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Prescan E., Pujic C., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Sciofione F.,
 RA Sekiguchi J., Sekowska A., Serot S.J., Serot P., Shin B.S., Soldo B.,
 RA Sorokin A., Taconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Yamakoshi A., Tanaka T., Tepstra P., Tognoni A.M.,
 RA Totsu T., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzengruber T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
 RA "The complete genome sequence of the Gram-positive bacterium *Bacillus*
 RT *subtilis*.";
 RL Nature 390:249-256(1997).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 CC (Probable).
 CC -1- SIMILARITY: Belongs to the nlpa lipoprotein family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X96983; CA65694.1; -;
 DR EMBL: U58859; AAB01345.1; -;
 DR EMBL: Z99108; CAB12739.1; -;
 DR FPIR: F69822; F69822.
 DR Subtilist; BG1588; yhcJ.
 DR InterPro: IPR004872; Lipoprotein_9.
 DR InterPro: IPR000437; Prok_lipoprot_9.
 DR Pfam: PF03180; Lipoprotein_9; 1.
 DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 KW Hypothetical protein; Membrane; Lipoprotein; Signal;

KM Complete proteome: Palmitate.
 FT SIGNAL 1 19
 FT CHAIN 20 263
 FT LIPID 20 20
 FT LIPID 20 20
 FT CONFLICT 79 79
 FT CONFLICT 118 118
 FT CONFLICT 170 170
 FT CONFLICT 259 260
 FT SEQUENCE 263 AA; 29186 MW; 0513C7D91CD2FBF3 CRC64;
 Query Match 21.4%; Score 296; DB 1; Length 263;
 Best Local Similarity 28.6%; Pred. No. 5e-13;
 Matches 78; Conservative 70; Mismatches 99; Indels 26; Gaps 9;
 QY 9 ICA--LASSIALAGCSNOSNEPALISKTAQTIKYGWAGPEQAVAEVAGVAKENYLT 66
 DB 6 ICSFVLVLVSFTACS---PSA---EHESIKIGI-AESDGAIMNYIAQZAEBA-GLD 54
 QY 67 VELVEFNDYAMNSAVSKGELDANMOMHPYLEKDSOEKGLNNLVIGNTFVPLAGYST 126
 DB 55 IOLIFPSDYAESDIALANKEDIDANAFOTISYOSTEKKYL-NLAPLDTTYITPMGIYSK 113
 QY 127 KIKTINELKDGATIAVPNDPSNLARALILEKQGIKLKDNLTLPSTLTDIVENPKKLV 186
 DB 114 RYERIRDISRGAVSVPCAFDFGRALTVLQAGLTLTKNGFNGTSGVDMIKDNPRHLKL 173
 QY 187 KEVDVSVARAIDVDVLAIVNNNVAQVGL-----TASENGVFEVDKSPYNNIVARAD 241
 DB 174 KAVRQO--DAVSGADVFEWKESEAKKAGLNPKKHTLKSGLMSEB---NNLIVRAE 226
 QY 242 NKDSKAIQDFVKAQYOTDEVEAEAKKQFKDGVIK 274
 DB 227 QDQREALQITLILYQADDPAAFIKEVYQGDIVR 259
 RESULT 16
 RBP2_PLAVB STANDARD; PRT: 2867 AA.
 AC Q00759; Q9NZM3;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Reticulocyte binding protein 2 precursor (PVRBP-2).
 GN RBP-2 OR RBP2.
 OS Plasmodium vivax (strain Belem).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_Taxid=31273;
 RN [1]
 RP SEQUENCE FROM N.A., AND REVISIONS TO 2438-2439.
 RX MEDLINE=20299192; PubMed=10838229;
 RA Galinski M.R., Xu M., Barnwell J.W.,
 RT "Plasmodium vivax reticulocyte binding protein-2 (PVRBP-2) shares
 RT structural features with PVRBP-1 and the Plasmodium yoelli 235 kDa
 RT rhoptry protein family." J. Mol. Biochem. Parasitol. 108:257-262(2000).
 RN [2]
 RP SEQUENCE OF 189-2439 FROM N.A.
 RX MEDLINE=92315338; PubMed=1617731;
 RA Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.,
 RT "A reticulocyte-binding protein complex of Plasmodium vivax
 RT merzoites." J. Cell Biol. 122:61-72(1992).
 RL Cell 63:1213-1226(1992).
 CC -1- FUNCTION: Involved in reticulocyte adhesion. Specifically binds to
 CC human reticulocyte cells.
 CC -1- SUBCELLULAR LOCATION: Membrane-bound (Probable).
 CC This SMS-S-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.
 CC -----
 DR EMBL: AF184623; AAF76525.1; -
 DR HSSP; P03069; IICM.
 KW Malaria; Receptor; Signal; Transmembrane; Repeat.
 FT SIGNAL 1 21
 FT CHAIN 22 2867
 FT DOMAIN 22 2805
 FT TRANSMEM 2806 2826
 FT DOMAIN 2827 2867
 FT DOMAIN 44 133
 FT DOMAIN 560 758
 FT DOMAIN 1112 1285
 FT DOMAIN 2758 2785
 FT REPEAT 2758 2761
 FT REPEAT 2762 2765
 FT REPEAT 2766 2769
 FT REPEAT 2770 2773
 FT REPEAT 2774 2777
 FT REPEAT 2778 2781
 FT REPEAT 2782 2785
 FT SEQUENCE 2867 AA; 331433 MW; 6E7D8CA71AFBFD3 CRC64;
 Query Match 8.1%; Score 112.5; DB 1; Length 2867;
 Best Local Similarity 22.7%; Pred. No. 18;
 Matches 65; Conservative 50; Mismatches 112; Indels 59; Gaps 13;
 QY 33 KTAQTI-----KYGWAGPEQAVAEVAGVAKENYLVLEVEFNDYAMNSAVS 83
 DB 1903 KNKAKTILGHVDTSAKYVGIKITPELALTELDD-ALTKAOLKEPESKNNVLEENMS 1961
 QY 84 K--GELDANMOMHPYL-----EKDSOEKGLNNLVIGNTF---VYPLAGYSTKI 128
 DB 1962 KNTNELDVHNGIDQAYKVVALEILLHSDEIDTKQDSKLIEMNQIYLVVLLNQYKNKI 2021
 QY 129 KTLNELKDGATIAVPN--DPSNLARA-----LILEKQGLIKLKNLTLPSTLD 176
 DB 2022 SSIKSEAEVSVKIGNSKKSELSKTCSDSYNDIALIEKO--TELQMRNSFTDEKT 2079
 QY 177 IVENPKLVKEVDTSVARAIDVDVLAIVNNNVAQVGL--TASENGVFEVDKSPYNN 234
 DB 2080 NTNSDSKLEKIKTDFESLKNALKTLF-----GEVNAIKASDNEHVQSKSE--VN 2129
 QY 235 IIVRADNKD-----SKAIQDFVKAQYOTDEVEAEAKKQFKDGVIK 274
 DB 2130 FALSEIKERTDIDSLTALDELKKGRTCEV--SRVKLKIDVYTK 2173
 RESULT 17
 GTF2_STRDO STANDARD; PRT: 1592 AA.
 ID GTF2_STRDO
 AC P27470;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Glucosyltransferase-1 precursor (EC 2.4.1.5) (GTF-I) (dextranucrase)
 DE (Sucrose 6-glucosyltransferase).
 OS Streptococcus downei (Streptococcus sobrinus).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_Taxid=1317;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=6715 / Serotype G;
 RX MEDLINE=91123227; PubMed=1704006;
 RA Abo H., Matsumura T., Kodama T., Ohta H., Fukui K., Kato K.,
 RA Kagawa H.,
 RT "Peptide sequences for sucrose splitting and glucan binding within
 RT Streptococcus sobrinus glucosyltransferase (water-insoluble glucan
 RT synthetase)." J. Bacteriol. 173:989-996(1991).
 RL J. Bacteriol. 173:989-996(1991).
 CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
 CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE

OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.

-1- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl)(N) = D-fructose + ((1,6)-alpha-D-glucosyl)(N+1).

-1- SUBCELLULAR LOCATION: Secretd.

-1- MISCELLANEOUS: GTP-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTP-SI SYNTHESIZES WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTP-SI SYNTHESIZES BOTH FORMS OF GLUCANS.

-1- SIMILARITY: Belongs to family 70 of glycosyl hydrolases.

-1- SIMILARITY: Contains 16 cell wall binding repeats.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: D90213; BAA14241.1; -

DR InterPro: IPR002479; CW binding.

DR InterPro: IPR003318; Glyco_hydro_70.

DR Pfam: PF01473; CW_binding_1.13.

DR Pfam: PF02324; Glyco_hydro_70; 1.

KW Transferase; Glycosyltransferase; Signal; Repeat; Dental caries.

FT SIGNAL 1 38 POTENTIAL.

FT CHAIN 39 1592 GLUCOSYLTRANSFERASE-I.

FT DOMAIN 39 1044 CATALYTIC (APPROXIMATE).

FT DOMAIN 1093 1592 GLUCAN-BINDING (APPROXIMATE).

FT REPEAT 1093 1592 6.5 X TANDEN REPEATS.

FT REPEAT 1158 1207 1.

FT REPEAT 1222 1272 2.

FT REPEAT 1287 1337 3.

FT REPEAT 1402 1451 4.

FT REPEAT 1514 1563 5.

FT REPEAT 1577 1592 6.

FT REPEAT 1592 1592 7 (INCOMPLETE).

FT SEQUENCE 1592 AA; 176167 MW; BCOA66D079351ECF CRC64;

Query Match 7.7%; Score 106; DB 1; Length 1592;

Best Local Similarity 26.6%; Pred. No. 24;

Matches 63; Conservative 29; Mismatches 89; Indels 56; Gaps 12;

QY 76 AMPNSAVSGELDANAMQKPYLER---DSQEKGLNLIYVGTFFVPLAGYSIKITL 131

DB 666 AMQNYOIGNGEILTSYRKGKALKOSDKGDATTRTSGVGVWNGNPNFSLDG---KVVAL 722

QY 132 N-----ELKDG-ATTAVPNDPSNLARALILEKOGILKLD-NTNLFSTT 174

DB 723 NMGALANQRYALAMSTQDQVATYATDADAS-----KAGLVKRTDENGILYFLN 772

QY 175 LDI-VENFKLVKEVDTVSAAALIDVDLAIVNNNTAGQVLT-----ASENGVFEV- 226

DB 773 DDLKGVANPQ--VSGFLQVWVPGAADDDIRVAASDSTADTKSLHODAAWDSRWMEG 830

QY 227 -----DKDSPYVNIIVA--RADNKDSKAIODFKAYQOTDEVEAKKQFQDVYK 274

DB 831 PSNFGSFATKEERYTNVAVIANNVKFKVSGITDFEMAPQ---YVSTGQGFQDSVYQ 884

RESULT 18

SCA4 RICAU STANDARD; PRT; 991 AA.

AC 09AUF4;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Antigenic heat-stable 120 kDa protein (Pst120) (120 kDa antigen)

DE (Protein ps 120) (Fragment).

GN SCA4 OR D.

OS Rickettsia australis.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;

Rickettsiaceae; Rickettsiae; Rickettsia.

NCBI_TaxId=787;

RM [1]

RP SEQUENCE FROM N.A.

RA Sekeyova Z., Roux V., Raoult D.

RT "Phylogenetic analysis of Rickettsia spp. by comparing sequence of the 'gene D' coding for an intracytoplasmic protein."

RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: AF187982; AAX31302.1; -

DR Antigen.

KW NON_TER 1 1

FT NON_TER 991 991

FT SEQUENCE 991 AA; 106518 MW; F2EDB94DDCD9749A CRC64;

Query Match 7.6%; Score 104.5; DB 1; Length 991;

Best Local Similarity 22.3%; Pred. No. 17;

Matches 61; Conservative 50; Mismatches 96; Indels 67; Gaps 14;

QY 27 EPAAISSTAQTIKVGMAGPEQVAEVAQVAKKYNLTVELV--FNDYAMP--NSAV 82

DB 682 EALASHKTNSTPTKISABIEVETGVAE-SITLAEDEKMTKGLVDGIYDKANPEYTKAV 740

QY 83 SKGELDANAMQKPYLERKDSQEKGLNLIYVGTFFVPLAGYSTKIKITNELKQAT-IA 141

DB 741 SRGVKSKTA---RP--EDKO-----AKDAASEVA 765

QY 142 VPNDPSNLARALILEKOGILKLDNTNLFSTLDIVENPKLVKEVDTVSAAALIDV 201

DB 766 LDRETONFTG--LKEONLEKPKPRDDIYNKAQDAELKAVITVLDNAPEKRVSE 822

QY 202 DL-----AVNV-----NNYAGVGLTASENGVEFEDKD---SPYNIIVARADNK 243

DB 823 EVMKKTSILINDISNLITEKVNFRAM--LSPSNLKTLEKKAETKVDYELVEKFGTK 880

QY 244 DS-KAIDPFKAYQOTDE--VEAPAKQFQDVYK 274

DB 881 SSTEBOOSFTKANLTDKTLKSKVRLQTIIDKLQ 914

RESULT 19

CH60 MYCGE STANDARD; PRT; 543 AA.

AC P47632; Q49358;

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE 60 kDa chaperonin (Protein Cpn60) (groEL protein).

GN GROEL OR GROEL OR MOPR OR MG392.

OS Mycoplasma genitalium.

OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.

NCBI_TaxId=2097;

RM [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 33530 / G-37;

RX MEDLINE=96026346; Pubmed=7569993;

RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A., Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M., Fritchman J.L., Weidman J.F., Small K.V., Sandhu M., Fritchman J.L., Nguyen D.T., Usterback T.R., Sauder D.M., Phillips C.A., Merrick J.M., Tomb J.-F., Dougherty B.A., Bort K.F., Hu P.-C., Lucier T.S., Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;

RA "The minimal gene complement of Mycoplasma genitalium.";

RL Science 270:397-403(1995).

[2]
 RN SEQUENCE OF 1-60 AND 466-543 FROM N.A.
 RP STRAIN=ATCC 33530 /G-37;
 RC MEDLINE=94075230; PubMed=8253680;
 RA Peterson S.N., Hu P.-C., Bolt K.F., Hutchison C.A. III;
 RT "A survey of the Mycoplasma genitalium genome by using random
 sequencing";
 RL J. Bacteriol. 175:7918-7930(1993).
 CC -1- FUNCTION: Prevents misfolding and promotes the refolding and
 proper assembly of unfolded polypeptides generated under stress
 conditions (By similarity).
 CC -1- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of
 7 subunits (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the chaperonin (HSP60) family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U39721; AAC71620.1; -;
 DR EMBL: U02252; AAD12515.1; ALT_INIT.
 DR EMBL: U02258; AAD12534.1; -;
 DR PIR: D64243; D64243.
 DR HSSP: P06139; 1GRU.
 DR TIGR: MG392; -;
 DR HAMAP: MF_00600; -; 1.
 DR InterPro: IPR001844; Chaperonin_Cpn60.
 DR InterPro: IPR002423; Cpn60_TCP-1.
 DR InterPro: IPR008950; GroEL-ATPase.
 DR Pfam: PF00118; Cpn60_TCP1; 1.
 DR PRINTS: PR00298; CHAPERONIN60.
 DR PRINTS: PR00304; TCOMPLEXTCP1.
 DR PROSITE: PS00296; CHAPERONINS_CPN60; 1.
 DR Chaperone: ATP-binding; Complete proteome.
 SQ SEQUENCE 543 AA; 58354 MW; 80FA5C7037C4BA8 CRC64;

Query Match 7.5%; Score 103.5; DB 1; Length 543;
 Best Local Similarity 20.8%; Pred. No. 9.3; Indels 103; Gaps 15;
 Matches 70; Conservative 55; Mismatches 109;

QY 3 FGR-----INGICALASGIAL-AGCSNQ-----SNEPAISK-----TAAQ 37
 DB 7 FGRDARTLLQGIINKANAVKTVGKQNVLEKFNAPLLTNCVGTAKIELSDPE 66
 QY 38 TIKVGMADPEQAVNAVAGQ-----VAKKKNTLVEIVENDYANPNSAVSKELDAN 90
 DB 67 NIGAKVISAVAVSTNDIADGTTATILAQEMTNIGIEI--NKGANPNV--IRRGIEPAS 123
 QY 91 AMOHKPYLEKDSQEKLANLVIVGNTFVYPLAGYSTKIKTLNELDGATIAVPNDPSN-- 148
 DB 124 LLIIE-LER-----YKIKITNIEIIOVAALISGSKEIGKL 159
 QY 149 LARALILEKQGLIKKNDNTLPSTLD-----IYENPK 183
 DB 160 IIAQAMALVKNQVIT--TTDAKNTINTLTETEGIEFGKTVASPYMVSDEKKEVVLQPKI 218
 QY 184 LV-----IKEDNSVVARAIDVDVLAVNNNVAGVLTSENGVFEDSDSYNNII 236
 DB 219 LVSSAKINIKIPLILBESVENGPNLIVADPFAEVTTTAAV-----KLRGTINAV 272
 QY 237 VARAD--NKDSKAIQDFYKAVQTDVEAKKQFED 270
 DB 273 AVKNEYGEROKALEDLAISSGTLAVNTEINSFED 309

AC P11001;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
 DE (Sucrose 6-glucosyltransferase).
 GN GTFI.
 OS Streptococcus downei (Streptococcus sobrius).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1317;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MF328;
 RX MEDLINE=87308014; PubMed=3040686;
 RA Ferretti J.J., Gilpin M.L., Russell R.R.B.;
 RT "Nucleotide sequence of a glucosyltransferase gene from Streptococcus
 sobrius MF28";
 RL J. Bacteriol. 169:4271-4278 (1987).
 CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS. THAT ARE THOUGHT
 TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
 OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
 AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
 CC -1- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl) (N) = D-
 fructose + ((1,6)-alpha-D-glucosyl) (N+1).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES BOTH
 WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
 FORMS OF GLUCANS.
 CC -1- SIMILARITY: Belongs to family 70 of glucosyl hydrolases.
 CC -1- SIMILARITY: Contains 19 cell wall binding repeats.
 CC -----
 DR EMBL: M17391; AAC63063.1; -;
 DR InterPro: IPR002479; CW_binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW_binding_1; 16.
 DR Pfam: PF0324; Glyco_hydro_70; 1.
 KW Transferrase; Glucosyltransferase; Signal; Repeat; Dental carrier.
 DB SIGNAL 1 38
 FT CHAIN 39 1597
 FT DOMAIN 39 1050
 FT DOMAIN 1099 1597
 FT DOMAIN 1099 1597
 FT REPEAT 1099 1132
 FT REPEAT 1163 1213
 FT REPEAT 1227 1277
 FT REPEAT 1292 1342
 FT REPEAT 1352 1399
 FT REPEAT 1406 1455
 FT REPEAT 1465 1512
 FT REPEAT 1519 1568
 FT REPEAT 1582 1597
 SQ SEQUENCE 1597 AA; 177080 MW; B9E86A200868798E CRC64;

Query Match 7.4%; Score 103; DB 1; Length 1597;
 Best Local Similarity 26.6%; Pred. No. 38;
 Matches 63; Conservative 28; Mismatches 90; Indels 56; Gaps 12;

QY 76 AMPSAVSKGELDANAMOHKPYLEK-----DSQKGNNTIVGNTFVYPLAGYSTIKTL 131
 DB 672 AMQVYQINGEILTSVYFGKALKQSDKQDATTRTSGVGVWGNQPNFSDG--KVAL 728
 QY 132 N-----ELKDG-ATIAVPNDPSNIALRILILEKQGLIKKND-NTNLFSTT 174

RESULT 20
 GTF1_STRDO STANDARD; PRT; 1597 AA.

Db 729 NMGAHANCERYALMSTVDGATVATDADAS-----KAGIVKRTDENGVLIFLN 778

Qy 175 LDI--VENPKQVYIKEVDTSVAPARAIDVDVLAVNNNAGVGLT-----ASENGVFE- 226

Db 779 DDLKGVANPQ--VSGFLQWVVPVGAADDIDIRVASTDSTGSKSIHQPAADSRVMEEG 836

Qy 227 -----DKDSRYVNTIYA-RADNKDSKAIDPFVAYQTDVEAEAKKQFQDGYIK 274

Db 837 FSNFQSPATKEEYITVIVANNVDFKFSWGITDFENAPQ---YVSTGQFIDSVIQ 890

RESULT 21

ACPEA_CLOAB STANDARD; PRT; 465 AA.

ID ACPEA_CLOAB

AC Q97K30;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Probable M18-family aminopeptidase 1 (EC 3.4.11.-).

GN ABEA OR CAC1091.

OS Clostridium acetobutylicum.

OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;

OC Clostridium.

OX NCBI_TaxID=1488;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;

RX MEDLINE=21359325; PubMed=1146286;

RA Noelting U., Breton G., Omechenko M.V., Makarova K.S., Zeng Q., Galsboon R., Lee H.W., Dubois J., Qiu D., Hitti J., Wolf Y.T., RA Tatusov R.L., Sabatche F., Doucet-Stamm L., Soucaille P., Daly M.J., RA Bennett G.N., Koonin E.V., Smith D.R.;

RT "genome sequence and comparative analysis of the solvent-producing bacterium Clostridium acetobutylicum."

RL J. Bacteriol. 183:4823-4838(2001).

CC -1- SIMILARITY: Belongs to peptidase family M18.

CC -1- COFACTOR: Zinc (by similarity).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC

DR EMBL: AE007625; AKK79065.1; -.

DR PIR: P97034; P97034.

DR HAMAP: MF_00466; - 1.

DR InterPro: IPR001948; Peptidase_M18.

DR Pfam: PF02127; Peptidase_M18; 1.

DR PRINTS: PRO0932; AMINO1PTASE.

KM Hypothetical protein; Hydrolase; Aminopeptidase; Metalloprotease;

KW Zinc; Complete proteome.

FT METAL 105 105 ZINC (POTENTIAL).

FT METAL 180 180 ZINC (POTENTIAL).

FT METAL 441 441 ZINC (POTENTIAL).

SO SEQUENCE 465 AA; 52011 MW; CB4C6D0A54C0A439 CRC64;

Query March 7.2%; Score 99.5; DB 1; Length 465;

Best Local Similarity 22.6%; Pred. No. 14;

Matches 52; Conservative 39; Mismatches 75; Indels 65; Gaps 11;

Qy 77 MPNSAVSKGBLDANAMQKRYLEKDSQKGLNVLVGNTFVPLAGYSTK-----IK 129

Db 1 MPNDLTK-----EYKNAMDK-YDDKQLKEVFALGDFKNFISNCTRECEYELK 50

Qy 130 TLNB-----LKDGATTAVPDPSPNLARALILEKQGL---IKLKONTNLF 171

Db 51 TAEKSGVRNEDILAKETLKEGDKYANRKGGL--IMFLIKELPLTYGFKLL-GAHD 107

Qy 172 STLDIVENPKVLVIKVDTSVAPARAIDVDVLAVNNNAG-----QVGLTASENGV 225

Db 108 SPRLDLKQNP-----LYEDTDLAMEETHYVGIGIKYQWTLPLAIGHVIV 152

Qy 226 EDKDSRYVNTIYARADNKDSKAIDPFVAYQTDVEAEAKKQFQDGYIK 275

Db 153 K-XDGTIVVCGVEDNDPVGVSILVHLASOLEKRAKSK-----VTEG 196

RESULT 22

SLAL_BACAA STANDARD; PRT; 814 AA.

ID SLAL_BACAA

AC P49051;

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE S-layer protein sap precursor (Surface layer protein) (Surface array protein).

GN SAP OR BA0895.

OS Bacillus anthracis (strain Ames), and

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=18094, 1392;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Ames;

RX MEDLINE=22608414; PubMed=12721629;

RA Read T.D., Peterson S.N., Tourasee N., Baillie L.W., Paulsen I.T., RA Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R., RA Holtzapple E.K., Osked O.A., Helgason E., Ristone J., Wu M., RA Kolonay J.F., Beanan M.J., Dodson R.J., Binkac L.M., Gwin M., RA Desoy R.T., Madun R., Daugherty S.C., Durkin A.S., Haft D.H., RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D., RA Benton J.L., Mamoud Y., Jjiang L., Hance I.R., Weidman W.C., RA Berry K.J., Platt R.D., Wolf A.M., Watkins K.L., Nierman W.C., RA Hazen A., Cline R., Redmond C., Thwaite J.E., White O., Salzberg S.L., RA Thomason B., Friedlander A.M., Koehler T.M., Hanna P.C., Kolsto A.-B., RA Fraser C.M.;

RT "The genome sequence of Bacillus anthracis Ames and comparison to closely related bacteria."

RL Nature 423:81-86(2003).

CC

RP SEQUENCE FROM N.A., AND SEQUENCE OF 30-40; 209-218; 281-291 AND 571-580.

RC STRAIN=Steine;

RX MEDLINE=95138020; PubMed=7836294;

RA Etienne-Toumelin I., Sirard J., Dufiot E., Mock M., Fouet A.;

RT "Characterization of the Bacillus anthracis S-layer: cloning and sequencing of the structural gene."

RL J. Bacteriol. 177:614-620(1995).

CC -1- FUNCTION: The S-layer is a paracrystalline mono-layered assembly of proteins which coat the surface of bacteria.

CC -1- SUBCELLULAR LOCATION: Cell wall.

CC -1- PTM: PROBABLY GLYCOSYLATED.

CC -1- SIMILARITY: Contains 3 S-layer homology (SLH) domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC

DR EMBL: AE017027; AAP24883.1; -.

DR EMBL: Z36946; CAAB5408.1; -.

DR PIR: I40048; I40048.

DR TIGR: BA0885; -.

DR InterPro: IPR003343; Big_2.

DR InterPro: IPR008964; Invasion_intimin.

DR InterPro: IPR001119; SLH.

DR Pfam: PF02368; Big_2; 1.

DR Pfam: PF00395; SLH_3.

KM Signal; Glycoprotein; Repeat; Cell wall; S-layer; Complete proteome.

FT SIGNAL 1 29

FT CHAIN 30 814 S-LAYER PROTEIN SAP.
 FT DOMAIN 33 93 SLH 1.
 FT DOMAIN 94 154 SLH 2.
 FT DOMAIN 155 213 SLH 3.
 SQ SEQUENCE 814 AA; 86620 MW; C1638D26A1C6B101 CRC64;

Query Match 7.2%; Score 99; DB 1; Length 814;
 Best Local Similarity 24.0%; Pred. No. 31;
 Matches 67; Conservative 40; Mismatches 120; Indels 52; Gaps 12;

QY 30 AISKTAACQIKYGVAGPEQA--VAEVAAGVAAKEKNLTYE-----LVEENDAMPASAKS 84
 DB 232 AVEKTKEDIKVTNKNANDKVLVEVT--LSEDKSAVELYSNLAQKQTVDVNKKVK 289
 QY 85 GILDANAMQKRELEKD-----SOEKLNNIVYNTFVYPLA----- 122
 DB 290 TEVAAGSLEAKTIEMADQTVVADEPTALQFTYKDNENGEVVSPEGIEFVTPAAEKINAKG 349
 QY 123 -----GISTKIKTNELKDQATTIVPNDPSNLARALILKQGLIKLKDNTNLFSTLLD 176
 DB 350 EITLAKGSTIVKAYK-KDGKVAESKEKVSAGAAVASISMTVAEONKADP-TSKD 407
 QY 177 IYENPKYVYKEVDPSVVARAIDVDLAVNNNVAQVGLTASENGVEVEDKSPYVNI 236
 DB 408 FKQNNK---VYEGDAVYQVELKDQPNVY----TGKVEYSINTEVAVDKATSKATVL 460
 QY 237 VA-----RADNDSKAIDPFYKAYQTVDEVEAKKQFQD 270
 DB 461 SAKGAPVKVTVKDSKG-KELVS--KTVEIEAPAKAME 496

RESULT 23
 ID HXA2_HAEIN STANDARD; PRT; 928 AA.
 AC P45354;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Heme/hemopexin-binding protein precursor (Heme:hemopexin utilization protein A).
 DE HxUA.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 22-35.
 RA STRAIN=DL42 / Serotype B;
 RC MEDLINE=9515556; PubMed=7815944;
 RA Cope L.D., Thomas S.E., Latimer J.L., Slaughter C.A.,
 RA Mueller-Eberhard U., Hansen E.J.;
 RT "The 100 kDa haem-hemopexin-binding protein of Haemophilus influenzae: structure and localization.";
 RL Mol. Microbiol. 13:863-873(1994).
 RN [2]
 RP SEQUENCE OF 1-30 FROM N.A.
 RA STRAIN=DL42 / Serotype B;
 RC MEDLINE=9570579; PubMed=7751272;
 RA Cope L.D., Yoger R., Mueller-Eberhard U., Hansen E.J.;
 RT "A gene cluster involved in the utilization of both free heme and heme:hemopexin by Haemophilus influenzae type b.";
 RL J. Bacteriol. 177:2644-2653(1995).
 CC -1- FUNCTION: BINDS HEME/HEMOPEXIN COMPLEXES.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

DR EMBL/ U08348; AAA74138.1; -.
 DR PIR; S54699; S54699.
 DR InterPro; IPR008638; Haemaggl. act.
 DR Pfam; PF05860; Haemaggl. act.; 1.
 DR Transprot; Signal; Repeat.
 FT SIGNAL 1 21
 FT CHAIN 22 928 HEME/HEMOPEXIN-BINDING PROTEIN.
 FT DOMAIN 101 679 6 X 6 AA APPROXIMATE REPEATS.
 FT REPEAT 101 106 1-1.
 FT REPEAT 205 210 1-2.
 FT REPEAT 279 284 1-3.
 FT REPEAT 410 415 1-3.
 FT REPEAT 635 640 1-4.
 FT REPEAT 674 679 1-5.
 FT DOMAIN 149 172 4 X 6 AA APPROXIMATE TANDEM REPEATS.
 FT REPEAT 149 154 2-1.
 FT REPEAT 155 160 2-2.
 FT REPEAT 161 166 2-3.
 FT REPEAT 167 172 2-4.
 SQ SEQUENCE 928 AA; 101228 MW; 67D4546A4B92390 CRC64;

Query Match 7.2%; Score 99; DB 1; Length 928;
 Best Local Similarity 23.7%; Pred. No. 36;
 Matches 62; Conservative 40; Mismatches 86; Indels 74; Gaps 13;

QY 5 KINGICALASGIALAGCSNOSNEPAISKTA-AQITKVGVMAGPEQA VAEVAQVAKEXY 63
 DB 119 EINVAAGLATTCKLERISENSNSYQFTRRTKDRQVLEKGLVADQGVAKS--QVINEG- 175
 QY 64 NLATVELVEFNDYMPN--SAVSKGELDANAMQKRELEKDSQSKGLNNIVYNTFVYPL 121
 DB 176 NITRAQ-----DFVANGDEVYINKNTN-----VEKNSTNGK-----VYLS 211
 QY 122 AGYSTKIKITNELKDQATTIAPNDPSNLARALILKQGLIKLKDNTNLFSTLLDIVENP 181
 DB 212 SGVN-----FTFTLPDGSISVA-----LEDN-----TVQGIKVE 241
 QY 182 KKLIVKVDTSVVAR--AIDVDLAVVNNNVAQVGLTASENGVEVEDKSPYVNI----- 234
 DB 242 GSIRAGEITLSAGRQALD---SLVMNNGVLEATKVSINKKGVVLSADVELNNESSNI 297
 QY 235 ---IIVRADNDSKAIDPFVK 253
 DB 298 KGEIVTFGADVTSNKLKDNK 319

RESULT 24
 ID TRME_CLOTE STANDARD; PRT; 459 AA.
 AC Q89952;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE tRNA modification GTPase trme.
 DE TRME OR THDF OR CTC00098.
 OS Clostridium tetani.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=Massachusetts / 588;
 RC MEDLINE=22457253; PubMed=12552129;
 RA Brueggemann H., Baumer S., Fricke W.F., Wierer A., Liesegang H.,
 RA Decker I., Herzberg C., Martinez-Arias R., Merkl R., Henne A.,
 RA Gottschalk G.;
 RT "The genome sequence of Clostridium tetani, the causative agent of tetanus disease.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321(2003).
 CC -1- FUNCTION: Exhibits a very high intrinsic GTPase hydrolysis rate. Involved in the biosynthesis of the hypermodified nucleoside 5-methylaminomethyl-2-thiouridine, which is found in the wobble position of some tRNAs (By similarity).

RL J. *Bacteriol.* 177:6874-6880(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-158 / JH642:
RX MEDLINE=99232519; PubMed=10216873;
RA Kapes R.W., Kempf B., Kneip S., Boon J., Gade J., Meier-Wagner J.,
RA Bremer E.;
RT "Two evolutionarily closely related ABC transporters mediate the
RT uptake of choline for synthesis of the osmoprotectant glycine betaine

in Bacillus subtilis":
Mol. Microbiol. 32:203-216(1999).
[3].

SEQUENCE FROM N.A.

RX MEDLINE=98044033; PubMed=9384377;
RZ Kunst F., Ogasawara N., Moser I., Albertini A.M., Alloni G.,
RZ Azavedo V., Bertero M.G., Bessieres P., Bolotin A., Borbeth S.,
RZ Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RZ Bouillere C., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RZ Choi S.K., Codan U.J., Connetton I.F., Cummings N.J., Daniel R.A.,
RZ Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RZ Enlian K.D., Errington J., Fabre C., Ferrari E., Foulger D.,
RZ Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RZ Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RZ Giuseppe G., Guy B.J., Haga K., Haelegh J., Harwood C.R., Henaut A.,
RZ Hilbert H., Holeppel S., Hosono S., Hulio M.F., Itaya M., Jones D.,
RZ Joris B., Katamata D., Kasahara Y., Klaert-Blanchard M., Klein C.,
RZ Kobayashi Y., Koester P., Koningsstein G., Krogh S., Kumano M.,
RZ Kurita K., Lapidis A., Lardinis S., Lauber J., Lazarevic V.,
RZ Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
RZ Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RZ Noone D., O'Reilly M., Ogawa K., Ogikawa A., Oudega B., Park S.H.,
RZ Parro V., Poh T.M., Portetelle D., Porwollik S., Prescott A.M.,
RZ Pressecan E., Puig P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RZ Rigser M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
RZ Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RZ Selighuchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RZ Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RZ Takeuchi M., Tanakaoshi A., Tanaka T., Terpetra P., Togroni A.,
RZ Tosato V., Uchiyama S., Vandenhof M., Vanter H., Vasarotti A.,
RZ Viari A., Wambut R., Wedler E., Wedler H., Weltzenegger T.,
RZ Witters P., Wigat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RZ Yoshida K., Yoshioka H.F., Zumschein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis".
RT Nature 390:249-256(1997).

- FUNCTION: MEMBER OF A HIGH AFFINITY MULTICOMPONENT BINDING-
PROTEIN-DEPENDENT TRANSPORT SYSTEM FOR CHOLINE.
SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
(Probable).

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).

EMBL, U38418; AA501534.1; -
DR EMBL, AF008930; AAC14358.1; -
DR EMBL, 299121; CAB15376.1; -
PIR: A69670; A69670.
Subtilisin; BG12635; OpABC.
Interpro: IPRO007210; OpABC.
Interpro: IPRO00437; Prok_lipprot_S.
Pfam: PF04069; Opnuc_1.
ProSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Transport; Amino-acid transport; Lipoprotein; Membrane; Signal;
Complete proteome; Palmitate.
SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 306 CHOLINE-BINDING PROTEIN.
FT LIPID 23 23 N-palmitoyl cysteine (Probable).
FT LIPID 23 23 S-dialcylglycerol cysteine (Probable).
FT VARIANT 8 8 L->W (IN STRAIN LH45).
FT VARIANT 13 13 A->S (IN STRAIN LH45).
FT VARIANT 32 32 A->S (IN STRAIN LH45).
FT VARIANT 84 84 D->E (IN STRAIN LH45).
FT VARIANT 111 111 G->R (IN STRAIN LH45).
FT VARIANT 146 146 E->K (IN STRAIN LH45).
FT VARIANT 155 155 T->N (IN STRAIN LH45).
FT VARIANT 194 194 G->S (IN STRAIN LH45).

[illegible]

```
CC -----
DR EMBL: X59771; CAA42442.1; -
DR PIR: S1530; FCSOAG.
DR InterPro: IPR004829; Surface antigen.
DR InterPro: IPR005877; GspF_Ysirk.
DR InterPro: IPR001893; Gram_pos_anchor.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR007756; RICH.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR Pfam: PF05062; RICH; 1.
DR Pfam: PF04650; Ysirk_signal; 1.
DR ProDom: PD153432; Surface_antigen; 1.
DR SMART: SM00409; Ig; 1.
DR TIGRfams: TIGR01167; LpxTG_anchor; 1.
DR TIGRfams: TIGR01168; Ysirk_signal; 1.
DR PROSITE: PS00847; GRAM_POS_ANCHORING; 1.
DR PROSITE: PS00835; IG_LIKE; FALSE NEG.
DR Cell wall; Peptidoglycan-anchor; Receptor; Signal;
KW Immunoglobulin domain.
SQ
FT SIGNAL 1 37
FT CHAIN 38 1135 IGA FC RECEPTOR.
FT PROPEP 1136 1164 REMOVED BY SORTASE (POTENTIAL).
FT DOMAIN 434 534 IG-LIKE.
FT DOMAIN 199 438 IGA-BINDING (POTENTIAL).
FT DOMAIN 439 826 IGA-BINDING (POTENTIAL).
FT DOMAIN 827 944 PRO-RICH.
FT SITE 1132 1136 LpxTG SORTING SIGNAL (POTENTIAL).
FT MOD_RES 1135 AMIDE-LINKED TO CELL WALL (POTENTIAL).
FT MOD_RES 1135 AMIDE-LINKED TO CELL WALL (POTENTIAL).
SQ
Query Match 7.0%; Score 97.5; DB 1; Length 1164;
Best Local Similarity 22.5%; Pred. No. 60;
Matches 67; Conservative 38; Mismatches 110; Indels 83; Gaps 14;
QY 8 GICALASGIALAGCSNQNPAISKTAQITKGVAMPEQAVAEVAGQVAKKXNTL- 66
DB 20 GAAVAVASLFGVSAHASE--LVKDDSVKTEVAKPPSAVQDQGNSSSLETTK 77
QY 67 -----VEIVFENDYAMPNSAVSGELDANAOMHKKPYLEDSOGEKGLNVLVYGN 115
DB 78 MEPTTIDIKAVPEPE--KTAGEISAITDTGKSEKOLQWKNKLNKD-----VDN 124
QY 116 TTVYPLAGYSTKIKTLNELKDGATIAVPDPSNLARALILEKQ-----GIKLDNTNL 170
DB 125 T-----ILSHKQ-----NEFK-----TKIDETNDSALLEENQENETNRLHKKH- 167
QY 171 FETTIDIVENPKKL-----VIKEVDTSVAAARAIDVDVLAVN-----NNVAGQVGLTASENG 223
DB 168 -----EEVEKDKKAAQOKTKLKSDTK-----VDLSNIDKELNHQKQVEKMAEQKGI 214
QY 224 FVEDKDSPYVNIIVAR-----ADNKDSKAIQ-----DFVKAAYQDVEVAEAKKQ 267
DB 215 TVEDDSMKIKIEDIRKQAOQADKEDAEVKVREBLGLFSTYKAGLDQEIQEHYKKE 272
RESULT 28
Y414 MYCGE STANDARD; PRT; 1036 AA.
AC P47653; P47654; Q49457;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MG414.
GN MG414.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=96026346; PubMed=7569993;
RA Frazer C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
```

```
RA Fleischmann R.D., Sult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fritchman J.L.,
RA Nguyen D.T., Uterback T.R., Saudak D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.,
RA "The minimal genome complement of Mycoplasma genitalium."
RA Science 270:1397-403 (1995).
RN [2]
RN SEQUENCE OF 52-146 AND 733-833 FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=94075230; PubMed=8253680;
RA Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III;
RA "A survey of the Mycoplasma genitalium genome by using random
RA sequencing."
RA J. Bacteriol. 175:7918-7930 (1993).
RL J. Bacteriol. 175:7918-7930 (1993).
CC -1- SIMILARITY: BELONGS TO THE MG414 / MG415 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL databank.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: U39723; AAC71641.1; -
DR EMBL: U01695; AAB01008.1; -
DR EMBL: U01804; AAD12330.1; -
DR PIR: H64245; H64245.
DR TIGR: MG414; -
DR Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 4 24
FT TRANSWEM 1004 1024 LEPT -> SRAS (TM REF. 2).
FT CONFLICT 733 736
FT SEQUENCE 1036 AA; 123179 MW; 12A21F00686A1A1 CRC64;
SQ
Query Match 7.0%; Score 97; DB 1; Length 1036;
Best Local Similarity 19.9%; Pred. No. 56;
Matches 62; Conservative 43; Mismatches 91; Indels 116; Gaps 13;
QY 71 EENDY-----AMPSAISKG---ELDANA-----MHKKPYLEDSOGEKGL- 107
DB 687 KENDYERKQVAFITKNSFKSKDNIDLINSLSFVNDPKKYPKRLPFTNEKKGIFFYL 746
QY 108 -----NNTL-----VIVGNFVYPLAGYSTKIKTL-----NEV- 134
DB 747 AEINNSNELFKYGSSESDSEIIDKQVFLSQNKPNRLRYLPKPHNTKLFVDKNDGFLY 806
QY 135 -KDGATIAVPN-----DPSLALALILEKQGL----- 161
DB 807 EKDKVPLVDNKLKIALDLKSSFKFLEDYQDLDFHFPSLINDQQLVLDLNTLSEKRLQ 866
QY 162 -----TKLDKNTNLFTTIDIVENPKKLVIKEVDTSVAAARAIDVDVLAV 206
DB 867 TTGNVAFNLKKEFINIHVFNKQFVLVFDVVRSKGLFIKGVNNDNQVFSI-SYDLKTT 925
QY 207 NNNVAGQVGLTASENGFEVDKDSPYVNIIVAR-----RADNKDSKAIQDF-----VKAYQ 256
DB 926 NNQTLTIVANGFDNSIWFDTISENQTLFKALSFYKNNLQFRVPDFNLKSODKSYE 985
QY 257 TDEVEA-EAKKQ 267
DB 986 VDKLEKNEIKKQ 997
RESULT 29
PUS3 OCEIH STANDARD; PRT; 339 AA.
AC Q8ES94;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Phosphoribosylformylglycinamidine cyclo-ligase (EC 6.3.3.1) (AIRS)
```

DE (Phosphoribosyl-aminimidazole synthetase) (AIR synthase).
GN P1UM OR 080747.
OS Oceanobacillus thelyensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.
OX NCBI_TaxID=182710;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HTE831 / DSM 14371 / JCM 11309;
RX MEDLINE=22220767; PubMed=1223376;
RA Takami H., Takaki Y., Uchiyama I.;
RT "Genome sequence of Oceanobacillus thelyensis isolated from the Iheya Ridge and its unexpected adaptive capabilities to extreme environments."
RT Nucleic Acids Res. 30:3927-3935(2002).
RL -1- CATALYTIC ACTIVITY: ATP + 2-(formamido)-N(1)-(5-phospho-D-ribose)1]acetamide = ADP + phosphate + 5-amino-1-(5-phospho-D-ribose)1]imidazole.
CC -1- PATHWAY: De novo purine biosynthesis; fifth step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC -1- SIMILARITY: Belongs to the AIR synthase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP004595; BAC12703.1; -;
DR HAVAP; MF_00741; -; 1.
DR InterPro; IPR000728; AIR_synth.
DR InterPro; IPR004733; PurM_c1igase.
DR Pfam; PF00586; AIRS; 1.
DR Pfam; PF02769; AIRS_C; 1.
DR TIGRfams; TIGR00878; purM; 1.
DR Purine biosynthesis; Ligase; Complete proteome.
KM Putine biosynthesis; 36326 MW; 99DDA067B2445F C6C64;
SQ SEQUENCE 339 AA; 36326 MW; 99DDA067B2445F C6C64;
Query Match 7.0%; Score 96.5; DB 1; Length 339;
Best Local Similarity 20.6%; Pred. No. 15;
Matches 52; Conservative 41; Mismatches 78; Indels 81; Gaps 11;
CY 26 NEPAIASTACTIVGVAAGPEQA-----VAEVAQVAKKKNL---TYELVE--- 71
DB 111 NDPSRI-----EAIVSIGAEQCEQAGALIGGETEMGMYDPPEYDLAGFVAGIVEKSA 165
CY 72 -----FNDYAMPNSAVSKGEI-----DANAMQHKPY 97
DB 166 MITGDKISGVVIVGLSSSGHSNYSILVKLIADVINGVPLGSLQVYDAVWAPRTIY 225
CY 98 -----LEKDSQKGLNNLVIVGNTFVYPLAGYSTIKTILNELDGATIAVPNDSNLAR 151
DB 226 AKSIQALKKXKVKLKGISH--ITG-----GGFDENIPRM--LPDGLCVLLLETMSWDIFE 274
CY 152 ALILEKQGLIKLKNNTMLFSTTLDIVENPKVLVKEVDTSVAARAIDVDVLAVNNNYA 211
DB 275 VHFHLEEKNTINREMGVFNWIGMA-----VVAEDVSIALQLEKVD-----EQA 323
CY 212 GQVGLTASNGV 223
DB 324 VVIKVTREEGV 335
RESULT 30
GLTI YEAST STANDARD; PRT; 2144 AA.
AC Q12680; Q12290;
DT 01-NOV-1997 (Rel. 35; Created)
DT 01-NOV-1997 (Rel. 35; Last sequence update)
DT 30-MAY-2000 (Rel. 39; Last annotation update)
DE Glutamate synthase [NADPH] precursor (EC 1.4.1.13) (NADPH-GOGAT).
GN GLTI OR YD1171C.

OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=CN36;
RX MEDLINE=97082505; PubMed=8923741;
RA Flierfeld P., Martegani M.P., Valenzuela L., Gonzalez A., Ballario P.;
RT "Sequence of the GLTI gene from Saccharomyces cerevisiae reveals the domain structure of yeast glutamate synthase."
RT Yeast 12:1359-1366(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c;
RA Pohl T.M.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: 2 L-glutamate + NADP(+) = L-glutamine + 2-oxoglutarate + NADPH.
CC -1- COFACTOR: Binds a 3Fe-4S cluster; FAD and FMN.
CC -1- SUBUNIT: Homotrimer.
CC -1- SIMILARITY: TO OTHER GLUTAMATE SYNTHASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X89221; CA61505.1; -;
DR EMBL; Z67750; CA61574.1; -;
DR EMBL; Z74219; CA68745.1; -;
DR GenOnline; 140414; -;
DR SGD; S0002330; GLTI.
DR GO; GO:0005623; C:cell; IDA.
DR GO; GO:0006537; P:glutamate biosynthesis; IEPI.
DR InterPro; IPR000785; Adnrdx_reductase.
DR InterPro; IPR002489; DUF14.
DR InterPro; IPR001327; FAD_pyr_redox.
DR InterPro; IPR006982; Glu_synth_centre.
DR InterPro; IPR006981; Glu_synth_NTN.
DR InterPro; IPR002932; Glu_synthase.
DR InterPro; IPR006005; Glu_synth_sub1.
DR InterPro; IPR001003; Pyridine_redox_2.
DR Pfam; PF04898; Glu_synth_central; 1.
DR Pfam; PF04897; Glu_synth_NTN; 1.
DR Pfam; PF01645; Glu_synthase; 1.
DR Pfam; PF01493; GKGXG; 1.
DR Pfam; PF00070; Pyr_redox; 1.
DR PRINTS; PR00419; ADKRDPTASE.
DR PRINTS; PR00368; FADNR.
DR PRINTS; PR00469; PNDKDTASEII.
DR TIGRfams; TIGR01317; GOGAT_sm_gam; 1.
KV Oxidoreductase; Iron-sulfur; 3Fe-4S; Flavoprotein; FAD; FMN; NADP; Glutamate biosynthesis; Zymogen.
FT PROPEP 1 53
FT CHAIN 54 2144
FT NP_BIND 1131 1183
FT METAL 1184 1184
FT METAL 1184 1184
FT METAL 1190 1190
FT METAL 1195 1195
FT CONFLICT 30 30
FT CONFLICT 166 172
FT CONFLICT 449 451
FT CONFLICT 1752 1752
FT CONFLICT 1752 1752
SQ SEQUENCE 2144 AA; 238200 MW; 5AA6A948EF995345 C6C64;
Query Match 7.0%; Score 96.5; DB 1; Length 2144;
Best Local Similarity 27.8%; Pred. No. 1.5e+02;
Matches 67; Conservative 24; Mismatches 89; Indels 61; Gaps 14;

QY 73 NDYAMNSAVSGELDANMAMQHKPYLEKDSOEKGLNNIVTGNTPYPLAGYSTKI-----128
 Db 1377 NDY-----VGKQ-LSGGIIVIKP--FKDSKFKSPSEN-VIVANTCFFYATSGTAFISGSA 1426
 QY 129 -KTALNELKQATIAVPNDPSNLA-----PALITLLEKQGLIKLMDNTN-----LFSITLL 175
 Db 1427 GEFGRGRNSGATIVIRIKGNNAFEYMGGRRAIVLSQMSLNAFGATGGLIYCLTSDYD 1486
 QY 176 DIVENPKLVKEVDTSVAAARAIIDVDLAIVNN-----NYAGQVGLTASENG-----222
 Db 1487 DFEVGIKIN-----DIVELESICDPVEIAFVKNLLOEHNNY-OSQDLAARILGNFNHYLK 1539
 QY 223 VFVEDKSPYNNIV--RADNKSKAIDDFKAVQT-----DDEVA-----BAKQF 268
 Db 1540 DFKVPTPYKVKVLKKEKKAAPAKAKATSEYTKKFRSQVEVDVNTILLANQFAKEOE 1599
 QY 269 K 269
 Db 1600 K 1600

RESULT 31
 VARTC HUMAN STANDARD; PRT; 382 AA.
 ID VARTC_HUMAN STANDARD; PRT; 382 AA.
 AC P21283;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Vacuolar ATP synthase subunit C (EC 3.6.3.14) (V-ATPase C subunit)
 GN ATP6V1C1 OR ATP6C OR V-ATP6C OR ATP6D.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Osteoclastoma;
 RA MEDLINE=94071935; PubMed=8250920;
 RA van Hille B., Vanek M., Richener H., Green J.R., Bilbe G.;
 RT "Cloning and tissue distribution of subunits C, D, and E of the human
 RL vacuolar H(+)-ATPase.";
 RL Biochem. Biophys. Res. Commun. 197;15-21(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=21574584; PubMed=11707601;
 RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plass C., Heinonen K.,
 RA Mrozek K., Sill H., Knuttila S., Kolitz J.E., Archer K.J.,
 RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.,
 RT "BALC, the human member of a novel mammalian neuroectoderm gene
 RL lineage, is implicated in hematopoiesis and acute leukemia.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98;13901-13906(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Bone marrow;
 RA MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshitsuki S., Carninci P., Prange C.,
 RA Rane S.S., Loquellano N.A., Peters G.J., Adamson R.D., Mullany S.C.,
 RA Bosch S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Hultyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Hultyk S.W.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99;16899-16903(2002).
 RN [4]
 RP SEQUENCE OF 36-382 FROM N.A.
 RC TISSUE=Brain
 RX MEDLINE=91056085; PubMed=2147024;
 RA Nelson H., Mandiyan S., Nouri T., Moriyama Y., Miedel M.C.,
 RA Nelson N.;
 RT "Molecular cloning of cDNA encoding the C subunit of H(+)-ATPase from
 RT bovine chromaffin granules.";
 RL J. Biol. Chem. 265;20390-20393(1990).
 CC -1- FUNCTION: Subunit of the peripheral V1 complex of vacuolar ATPase.
 CC Subunit C is necessary for the assembly of the catalytic sector of
 CC the enzyme and is likely to have a specific function in its
 CC catalytic activity. V-ATPase is responsible for acidifying a
 CC variety of intracellular compartments in eukaryotic cells.
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
 CC H(+) (Out).
 CC -1- SUBUNIT: V-ATPase is a heteromultimeric enzyme composed of a
 CC peripheral catalytic V1 complex (components A to H) attached to
 CC an integral membrane V0 proton pore complex (components a, c, c',
 CC c'', and d).
 CC -1- SIMILARITY: Belongs to the V-ATPase C subunit family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL; X69151; CAA48903.1; -;
 DR EMBL; AF363578; AAL50383.1; -;
 DR EMBL; BC010960; AAL10960.1; -;
 DR EMBL; J05682; AAA36803.1; -;
 DR PIR; JN0907; JN0907.
 DR Genew; HGNC:856; ATP6V1C1.
 DR MM; 603097; -;
 DR GO; GO:0016669; C:proton-transporting two-sector ATPase complex; TAS.
 DR GO; GO:0046961; F:hydrogen-transporting ATPase activity, rota. .; TAS.
 DR GO; GO:0015992; P:proton transport; TAS.
 DR InterPro; IPR004907; V-ATPase_C.
 DR Pfam; PF03223; V-ATPase_C; 1.
 DR Hydrolase; ATP synthase; Hydrogen ion transport.
 KM
 SQ SEQUENCE 382 AA; 43941 MW; 5626E2B2B2D66BA7 CRC64;

Query Match 6.9%; Score 96; DB 1; Length 382;
 Best Local Similarity 23.0%; Pred. No. 19;
 Matches 70; Conservative 47; Mismatches 99; Indels 88; Gaps 16;

QY 18 LAGCSNNSPAPAKSTAAQTIKVG---VMAGPEQAVAEVAGVAKENLTVELVEFND 74
 Db 22 LHAATSKNNNAVATSKFNIPDLKVTGLDVLVG-----LSDELAKLD- 62
 QY 75 YAMPNSAVSGKGLDANMAMQHKPYLEKDSOEKGLNNIVTGN-----NTFYPLAGYSTK 127
 Db 63 -AFVEGVVK-----VAQYMDVIVEDSKDQKQENILANGVDLVYIIRFGQDMAKYPIK 115
 QY 128 -IKTLNELKQATIAVPNDPSNLA-----PALITLLEKQGLIKLMDNTNFS-TTLDIVNPKKL 184
 Db 116 QSLKNSITELIANGVTQIDDLKRSASATNLT--KNTLONLEKQVAGSLITSLAE----- 168
 QY 185 VIKEDVTSVAAARAIIDVDLAIV-----NNNYAGQVGLTAS-----ENGVFEDKSPYNN 234
 Db 169 IYKQDVFVLDSEYL--VTLLVVPPLTNNDWI KQYETLAEMVVPSSVVLSDQDSYLCN 226
 QY 235 IIVARAADKDSKALIDPF-----VKAYOTDEVEEA-----KKQFKGV 272
 Db 227 VTLFR-----KAVDFHKAENKFTIVRDQYEEKAKDEENRSLTDKKQF--GP 278
 QY 273 IKGM 276

Db 279 LVFW 282

RESULT 32

ARO_A_CAMTJ STANDARD; PRT; 428 AA.

AC P52312; Q9P36; 01-OCT-1996 (Rel. 34, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPs).

GN AROA OR Cj0895C.

OS Campylobacter jejuni.

OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteres; Campylobacteraceae; Campylobacter.

OX NCBI_TaxId=197;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN=81116;

RX MEDLINE=97128776; PubMed=8973316;

RA Woesten M.M.S.M., Dubbink V.H.J., van der Zeijst B.A.M.;

RT "The *aroA* gene of *Campylobacter jejuni*."

RL Gene 181:109-112(1996).

RN (2)

RP SEQUENCE FROM N.A.

RC STRAIN=NCIC 11168;

RX MEDLINE=20150912; PubMed=10688204;

RA Parthill J., Wren B.W., Mungall K., Ketley J.M., Churcher C., Basham D., Chillingworth T., Davies R.M., Felwell T., Holroyd S., Jagels K., Kallishnev A.V., Moule S., Pallen M.J., Penn C.W., Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M., Whitehead S., Barrett B.G.;

RT "The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences."

RL Nature 403:665-668(2000)

CC -1- CARBONYL ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate = phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.

CC -1- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway; sixth step.

CC -1- SUBUNIT: Monomer (By similarity).

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).

CC -1- SIMILARITY: Belongs to the EPSP synthase family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL, X89371; CAA61554.1; -

CC EMBL, AL139076; CAB73153.1; -

CC PIR, G81362; G81362.

CC PIR, JCS338; JCS338.

CC HAMAP, MF_00210; -; 1.

CC InterPro, IPR006264; AROA.

CC InterPro, IPR001986; EPSP_synth.

CC Pfam, PF00275; EPSP_synthase; 1.

CC ProDom, PD001867; EPSP_synthase; 1.

CC TrEMBL, TIGR01356; *aroA*; 1.

CC PROSITE, PS00104; EPSP SYNTHASE 1; 1.

CC PROSITE, PS00885; EPSP SYNTHASE 2; 1.

CC Aromatic amino acid biosynthesis; transferase; Complete proteome.

CC CONFLICT 152 152 K -> N (IN REF. 1).

CC CONFLICT 158 158 F -> Y (IN REF. 1).

CC CONFLICT 179 179 D -> N (IN REF. 1).

CC CONFLICT 183 183 T -> A (IN REF. 1).

CC CONFLICT 217 217 N -> S (IN REF. 1).

CC CONFLICT 241 241 V -> A (IN REF. 1).

CC SCIENTIFIC 428 279 LVFW 282

Query Match 6.9%; Score 96; DB 1; Length 428;

Best Local Similarity 22.1%; Pred. No. 22;

Matches 62; Conservative 54; Mismatches 103; Indels 62; Gaps 14;

Db 8 GICALAGG-IALAGCSQNSNEP-AAIKSTAAQITIKVGMAGPEQVAEVAQVAKKYNL 65

99 GFLAGISGFVLSGDKLNRRPKRISKPLTQ---IG-----ARIQRREANLAPL 146

Qy 66 TVE---LVEENDYAMPNSAVSKGELDANMQ-----HKPYEKSGQ---KGLNNLVI 112

Db 147 CIEGQKLKAFNFSEISSAQVKTAMILSAFADNVCTFSEISIRHNSNNMLKAMKAPIR 206

Qy 113 VGNFTFVPLAGYSTKTKTLNELKDGATIAVPNDPSN---LARAALLLEKQGLIKLNDT 168

Db 207 VUSD-----GLSRIPLKKPLKAQNTIIPNDPSAFYVLAAILDPSQITLK----- 255

Qy 169 NLESTTLDIYENPKKLVKEVDTSVARAIDVDVLAIVNNV---AQVGLTASE-NGVFFV 225

Db 256 -----NILNLPRIEAYKILQMGAK---LEMTITQNDFTETIGIRVSSKNGIEV 304

Qy 226 EDK-----DSPYNTIVARADNKDKAIDPFVKAQOTDEV 260

Db 305 KDNIMLIDEPALALAFALAKGKSSLINAKELRVESDRI 345

RESULT 33

SPI_RARPA STANDARD; PRT; 525 AA.

AC Q05308; 01-FEB-1995 (Rel. 31, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Serine protease I precursor (EC 3.4.21.-) (Rpi).

OS Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Micrococciaceae; Karobacteraceae; Karobacter.

OX NCBI_TaxId=13243;

RN (1)

RP SEQUENCE FROM N.A., AND SEQUENCE OF 212-222; 224-238 AND 240-244.

RC STRAIN=YLM-50;

RX MEDLINE=93093426; PubMed=1339445;

RA Shimoi H., Iimura Y., Obata T., Tadenuma M.;

RT "Molecular structure of Karobacter faecitabidus protease I. A yeast-lytic serine protease having mannose-binding activity."

RL J. Biol. Chem. 267:25189-25195(1992).

RN (2)

RP SEQUENCE OF 212-247.

RX MEDLINE=92138668; PubMed=1778983;

RA Shimoi H., Tadenuma M.;

RT "Characterization of Karobacter faecitabidus protease I, a yeast-lytic serine protease having mannose-binding activity."

RL J. Biochem. 110:608-613(1991).

CC -1- FUNCTION: MAJOR SERINE PROTEASE EXHIBITING LYTIC ACTIVITY TOWARD LIVING YEAST CELLS. SIMILAR TO ELASTASE IN ITS SUBSTRATE SPECIFICITY AND HAS A LECTIN-LIKE APPENDIX FOR MANNOSE. MANNOPROTEIN MAY BE THE NATIVE SUBSTRATE FOR RPI.

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- SIMILARITY: Contains 1 ricin B-type lectin domain.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL, D10753; BA001585.1; -

CC PIR, A45053; A45053.

CC HSP, P00778; IGSA.

DR InterPro: IPR004236; AL protease.
DR InterPro: IPR009003; Cys_Ser trypsin.
DR InterPro: IPR001254; peptidase_S1.
DR InterPro: IPR001316; peptidase_S1.
DR InterPro: IPR000772; Ricin_B_lectin.
DR InterPro: IPR008997; Ricin_B_lectin.
DR Pfam: PF02983; AL protease; 1.
DR Pfam: PF00652; Ricin_B_lectin; 3.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00865; ALTYCPTASE.
DR SMART: SM00458; RICIN; 1.
DR PROSITE: PS00134; TRYPsin_HIS; 1.
DR PROSITE: PS00135; TRYPsin_SER; 1.
DR PROSITE: PS50231; Ricin_B_lectin; 1.
DR KEGG: K00001; Serine protease; Mannose-binding; signal; Zymogen;
KW Lectin.
FT SIGNAL 1 32
FT PROPEP 1 32
FT CHAIN 212 211
FT DOMAIN 212 211
FT DOMAIN 356 355
FT DOMAIN 401 525
FT DISULFID 223 239
FT DISULFID 310 330
FT DISULFID 346 376
FT DISULFID 412 431
FT DISULFID 453 472
FT ACT SITE 238 238
FT ACT SITE 270 270
FT ACT SITE 352 352
SQ SEQUENCE 525 AA; 55654 MW; DA3BCFD330EB61 CRC64;
Query Match 6.9%; Score 96; DB 1; Length 525;
Best Local Similarity 24.2%; Pred. No. 28; Indels 88; Gaps 14;
Matches 75; Conservative 30; Mismatches 117;
11 ALASGIALAGCNSQNSPAAIKTAQITKVG-VNAGEQAVAEVAGQVAKENLVTEL 69
8 ALFSAIALVGAIGASVGLAASANSASPAAATVQASSSATSVAAISKQDGEVLAI 67
70 VERNDYAMNSAVSKG-ELDANAMQHPYLEKDSQEK--GL-----NNLVYGNTF----- 117
68 V--RD.LKTKTQAKRIKLEKRAQLEPRLQKKLKKKFAGLWISKNGKIVGVTTKRA 125
118 -YTPLAGSTK--KTLNELKDGATLAVPNDPSNLARLILKEKGLIKLKNLTFLSTT 174
126 KVVKKAGATPKIVKSNLTLTKRAIKISGNAPS-----DIXVNSW 167
175 LDIENPKKLVIXEVDTSVAPARAIDVDLAVNNNVAQVGLTASENGVFVEDKDSPYN 234
168 VDPATN--KVVI--EASKKAAKAA-----ATAAGLTAGTYEITVSD-----D 206
235 IYARADNKSRAIQDFVAK-----QDVEAEAKK 266
207 VIVPVDYWGDAISGCTLAFPYGGFLTAGCAVEKGHLIKTMTGGQIGTVEA---S 263
267 QFKDGVIKGM 276
264 QFGDGI:DAW 273

OC Pseudomonadaceae; Azotobacter.
OX NCBI_TaxID=354;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UMR;
RA "an active group II intron has invaded the genus Azotobacter and is
RT inserted within the essential groEL gene";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Prevents misfolding and promotes the refolding and
CC proper assembly of unfolded polypeptides generated under stress
CC conditions (By similarity).
CC -1- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of
CC 7 subunits (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the chaperonin (HSP60) family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC EMBL: AY057439; AL25964.1; -.
DR HAMAP: MF_00600; -1.
DR InterPro: IPR001844; Chaperonin Cpn60.
DR InterPro: IPR002423; Cpn60/TCF-1.
DR InterPro: IPR008950; GroEL-ATPase.
DR Pfam: PF00118; Cpn60_TCF; 1.
DR PRINTS: PR00298; CHAPERONIN60.
DR PRINTS: PR00304; TCOMPLEXTCP.
DR PROSITE: PS00296; CHAPERONIN60; 1.
DR Chaperone; ATP-binding.
KW SEQUENCE 546 AA; 56875 MW; 8E1F4C93E42A619 CRC64;
Query Match 6.9%; Score 96; DB 1; Length 546;
Best Local Similarity 20.8%; Pred. No. 30;
Matches 74; Conservative 49; Mismatches 102; Indels 130; Gaps 16;
8 GICMALSQ-----IALAGCNSQNSPAAIKTAQ--TIKGVNAGEQAV 51
103 GLAVAAAGNPMPLDKRGIDKATTAIVAEKSLAKPESDSKALNOVTTIANSDESIGNII 162
52 AEVAGVAKENLVTE-----LVY-----FNDYAMPNSAVSKGELD----- 88
163 AEMNNVYKGV-ITYBEGSGLENELSVYEGMDFRGYLSPIFNKPDVTVAELDNPILL 221
89 -----ANAMQHPYLE-----KDSQEGANNVIVGNTFYVPLA-----GY 124
222 LVDKKISNIRELPLVEAVAKSGRPLLIVAEDVEGALTLVNNMRGIVKVAAYAPGF 281
125 STKIKTNEKDGATLAVPNDPSNLARLILKEKGLIKLKNLTFLSTLDIENPKKL 184
282 NDRKAMLVY---LAI-----LTGATVISKVL-----SLESATLHLGKPKAL 323
185 VIKEDTSV--AAPAIDVDLAVNNNVAQVGLTASENGVFVEDKDSPYNIIYARADN 242
324 VLNKENTTINHGAGAQDIEAAVAQIRKQ-----IEFTSSPY----- 360
243 KDSRAIQD-----FYKAVQDVEAEAKK-----QFKDGVIKG 275
361 -DRKCOERLAKAGVAVIKVGAATVEKKEKAAVEALAHATRAAVEGVVPG 414

RESULT 34
ID CH60_AZOVI STANDARD; PRT; 546 AA.
AC O8GBA4;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 60 kDa chaperonin (Protein Cpn60) (groEL protein).
OS GROEL OR GROEL.
GN Azotobacter vinelandii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

RESULT 35
ID RRP5_YEAST STANDARD; PRT; 1729 AA.
AC O05022;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE rRNA biogenesis protein RRP5.
GN RRP5 OR FM11 OR YMR229C OR YMR959.11C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=5288C / AB972;
RC MEDLINE=97313268; PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dehman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagsis K., Iye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA Rice P., Skellern J., Walsh S., Whitehead S., Barrall B.G.,
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
XIII."
RL Nature 387:90-93(1997).
RM [2]
RP CHARACTERIZATION.
RX MEDLINE=97051828; PubMed=8896463;
RA Venema J., Tollervey D.,
RT "RRP5 is required for formation of both 18S and 5.8S rRNA in yeast."
RL EMBO J. 15:5701-5714(1996).
CC -!- FUNCTION: INVOLVED IN THE BIOGENESIS OF RRNA. REQUIRED FOR THE
CC FORMATION OF 18S AND 5.8S RRNA.
CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar.
CC -!- SIMILARITY: Contains 4 HAT repeats.
CC -!- SIMILARITY: Contains 11 SI motif domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; 249939; CA90200.1; -.
DR PIR; S57596; S57596.
DR HSSP; P05055; 1SR0.
DR Germline; 142904; -.
DR SGD; S0004842; RRP5.
DR GO; GO:0005731; C:nucleolus organizer complex; IPI.
DR GO; GO:0030515; P:snRNA binding; IPI.
DR GO; GO:0030490; P:processing of 20S pre-rRNA; IPI.
DR InterPro; IPR003107; HAT.
DR InterPro; IPR008984; Nucleic_acid_OB.
DR InterPro; IPR003029; SI.
DR InterPro; IPR008941; TPR-like.
DR Pfam; PF00575; SI_7.
DR SMART; SM00386; HAT; 6.
DR SMART; SM00316; SI; 12.
DR PROSITE; PS50126; SI; 11.
KM Nuclear protein; rRNA processing; Repeat.
FT DOMAIN 119 200 SI MOTIF 1.
FT DOMAIN 338 410 SI MOTIF 2.
FT DOMAIN 510 580 SI MOTIF 3.
FT DOMAIN 607 676 SI MOTIF 4.
FT DOMAIN 690 769 SI MOTIF 5.
FT DOMAIN 794 863 SI MOTIF 6.
FT DOMAIN 885 971 SI MOTIF 7.
FT DOMAIN 1003 1083 SI MOTIF 8.
FT DOMAIN 1088 1159 SI MOTIF 9.
FT DOMAIN 1177 1245 SI MOTIF 10.
FT DOMAIN 1265 1336 SI MOTIF 11.
FT REPEAT 1455 1487 HAT 1.
FT REPEAT 1561 1594 HAT 2.
FT REPEAT 1632 1664 HAT 3.
FT REPEAT 1666 1701 HAT 4.
FT SEQUENCE 1729 AA; 193133 MW; 39BF46E55873B0A CRC64;

Query Match 6.9%; Score 96; DB 1; Length 1729;
Best Local Similarity 21.0%; Pred No 1 2e-02.
Matches 57; Conservative 46; Mismatches 82; Indels 86; Gaps 14;
QY 77 MPNSAVK-----GE-----LDANAMQH-----KPYLRKDSQEKIANLVI 112
DB 634 LPNSEISEVFPRKEEHLRLGQYIVLVLDVDDRRRIINATCKSNQNAQKQNTIENIV 693
QY 113 VGNFTVFPPLAGYSTKIKITLBELKDGATIANVN-----DPSNARAL 153
DB 694 PGRITL-----TVHVEIKTQDSVVEIPDVGRLGVIYVGHLSDSRIEONRAQLKLR 745
QY 154 ILLEKQGLIKKD-NTLFFSTTL--DIVENPKLVI-----KEYDTSVABAID 199
DB 746 IGTLTGLVIDKDRTRTVFMNLSKSLIKDAKKTLPITDVKDLNKDVPMAIYKIS 805
QY 200 DVDLAIVNNNNAAG-VGLTASENGVEVDKD--SPYVN---IIVAPADNKSALQDP 251
DB 806 DKGLFVAFN--GKFIGLTVPSYAVDSRDIDISKAFYINQSVYLLRTDCKOK---F 858
QY 252 VKAYQTEVEAEAKQK-----DGVIKGM 276
DB 859 LLSKAPKVEEKKEKVESNIEDPVSSIKSW 889
RESULT 36
AROA_AQJAE STANDARD; PRT; 431 AA.
ID AROA_AQJAE
AC 067494;
DT 30-MAY-2000 (Rel. 39, Created)
DR 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE enolpyruvylshikimate-3-phosphate synthase) (EPPS synthase) (EPPS).
GN AROA_Q01536.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
CC NCB1_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Snead M.A., Keller M., Aulay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.,
RT "The complete genome of the hyperthermophilic bacterium Aquifex
aeolicus."
RL Nature 397:353-358(1998).
CC -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
CC sixth step.
CC -!- SUBUNIT: Monomer (by similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- SIMILARITY: Belongs to the EPPS synthase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000744; AAC07443.1; -.
DR PIR; D70433; D70433.
DR HAMAP; MF_00210; -; 1.
DR InterPro; IPR006264; AROA.
DR InterPro; IPR001986; EPPS synth.
DR Pfam; PF00275; EPPS synthase; 1.
DR ProDom; PD001867; EPPS synthase; 1.
DR TIGRFAMs; TIGR01356; aroA; 1.
DR PROSITE; PS00104; EPPS SYNTHASE 1; 1.
DR PROSITE; PS00885; EPPS SYNTHASE 2; 1.
KW Aromatic amino acid biosynthesis; Transferase; Complete proteome

SO SEQUENCE 431 AA; 47793 MW; EF842512EBE41D2A CRC64;
Query Match 6.9%; Score 95.5; DB 1; Length 431;
Best Local Similarity 24.9%; Pred. No. 24;
Matches 58; Conservative 33; Mismatches 95; Indels 47; Gaps 12;
76 AMNSAVSNGELDANAMQH--KPYLEKDSQEKLNVLVVGNTFVYLP-LAGSTKIKTLN 132
165 AQYKSLHLAGLAEGETVEVEPYLSRDTERRKLK--LGAAYITTPERRGHIVIKSGQ 222
133 ELKQGTIAVPNDPSNLA--RALILLEKQGLIKLKDNTNLFTSTLD-----I 177
223 ELQ-GTEVCPADPSAAVFAALATLAPGEIRLKE--VLNPTRGDFYRLKLENGDIS 279
178 VEPK-----LVKEVDTSVAA-----IDVVDLAVNNVAGVGLTASN 221
280 FENYRELSEPMADLVLPVNLKPKVSPPEVEPTLIDITPLAVMAFAD--GVSEVAG 337
222 GVFEVDKDSFYNNIIVARADNKSKAIQDFVAVQTEVEAARQKQFQGVIK 274
338 AKELARKESDRIRKAITNL-RKLGQVEEF-----EDGALHGTGKIGGVIE 384
RESULT 37
FLD_HELPY STANDARD; PRT; 684 AA.
ID FLD_HELPY
AC P96786;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Flagellar hook-associated protein 2 (HAP2) (Filament cap protein)
DE (Flagellar cap protein).
GN FLD OR HP0752.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.; SEQUENCE OF 1-15, AND CHARACTERIZATION.
RC STRAIN=KCTC 0217BP;
RX MEDLINE=20026809; Pubmed=10559162;
RA Kim J.S., Chang J.H., Chung S.I., Yum J.S.;
RT "Molecular cloning and characterization of the Helicobacter pylori
J. Bacteriol. 181:6969-6976(1999)."
RL [2]
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; Pubmed=9252185;
RA Tomb J.F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klein H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus K., Richardson D., Dodson R., Khaliq H.G., Glodek A.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Meldrum J.M., Fujii C., Bowman C., Matthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
pylori.";
RT Nature 388:539-547(1997).
CC -1- FUNCTION: REQUIRED FOR THE MORPHOGENESIS AND FOR THE ELONGATION OF
THE FLAGELLAR FILAMENT BY FACILITATING POLYMERIZATION OF THE
FLAGELLIN MONOMERS AT THE TIP OF GROWING FILAMENT. FORMS A CAPPING
STRUCTURE, WHICH PREVENTS FLAGELLIN SUBUNITS (TRANSPORTED THROUGH
THE CENTRAL CHANNEL OF THE FLAGELLUM) FROM LEAKING OUT WITHOUT
POLYMERIZATION AT THE DISTAL END. ESSENTIAL TO COLONIZE AND
ESTABLISH INFECTION IN GASTRIC MUCOSA AS A RESULT OF ITS ESSENTIAL
ROLE IN MOTILITY. HAS EFFECT ON PLA GENE TRANSCRIPTION.
CC -1- SUBUNIT: Homopentamer (by similarity).
CC -1- SUBCELLULAR LOCATION: Flagellar.
CC -1- SIMILARITY: Belongs to the fld family.

CC This SWISS-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
DR EMBL; U82981; AAC3599.1; --
DR EMBL; AE000587; AAD07800.1; --
DR PIR; H64613; H64613.
DR TIGR; HP0752; --
DR InterPro; IPR003481; FLD.
DR Pfam; PF02465; FLD; 1.
KW Flagellum; Coiled coil; Complete proteome.
FT INIT MET 0
FT DOMAIN 617 646
FT VARIANT 58 58
FT VARIANT 85 85
FT VARIANT 90 90
FT VARIANT 192 192
FT VARIANT 230 230
FT VARIANT 257 257
FT VARIANT 269 269
FT VARIANT 362 362
FT VARIANT 367 367
FT VARIANT 372 372
FT VARIANT 421 421
FT CONFLICT 653 684
SO SEQUENCE 684 AA; 74000 MW; 4CAA3965425C4E0 CRC64;
Query Match 6.9%; Score 95.5; DB 1; Length 684;
Best Local Similarity 19.4%; Pred. No. 42;
Matches 60; Conservative 53; Mismatches 118; Indels 79; Gaps 14;
4 GKINGICAL-ASGIALAGCSNENPAISKTAQTI--KVGVA-----GPEQA 50
329 GKLEGVVSLNGQKLDLSLTKRESYTSSENTDAIOLNAKESAPRANEGKIVNSKTG 388
51 VAEVAGVAKKKNLTVELVEFNDYAMPNSAVSKGELDANAMQHPYLEKDSQEKLNVL 110
389 MLTIKGDALGKASL-----KDLGLNAGMVOSYENSQVTLFMSKRLQVAD----- 434
111 VIVNTFVYPLAGSTKIKTLNLEKQATIAV-----PNDPSNLAAALLLEKGLIKK 165
435 ---SAFTYGVSTIRPTNEVNDVISGVNITLTQTEPNKP-----ALISVS-----R 478
166 DNTNLFTSTLDIVENPKLVIK-----EVDTSVAA--SAIDVDL--AVNNNNVAGV 214
479 DNGAIIISLTFEYVANNELPKLDEDRYDADTKIKGIFGVGDIPAIRSLNNVFSYV 538
215 -----GLTASENGVFPEDKSPYNNIIVARADNKSKAIQDFVAVQTEVEA 262
539 HTDNGVESLMTKYGLSLDDKGVMSLDEAK-----LSSALNSNPAQVDFP--VGSOSKDM 590
QY 263 EAKKQFQGV 272
Db 591 GGRRIHQEG 600
RESULT 38
SCA4_RICRH STANDARD; PRT; 1013 AA.
ID SCA4_RICRH
AC Q9A081;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Antigenic heat-stable 120 kDa protein (Psl20) (120 kDa antigen)
DE (Protein PS 120) (Fragment).
CN SCA4 OR D
OS Rickettsia rhipicephali.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;

CC Rickettsiaceae; Rickettsiaceae; Rickettsia.
OX NCBI_TaxID=33392;
RN [1]
RP SEQUENCE FROM N.A.
RA Sekoyeva Z., Roux V., Raoult D.;
RT "Phylogenetic analysis of Rickettsia spp. by comparing sequence of the
RL 'gene D' coding for an intracytoplasmic protein."
RL Submitted (May-1999) to the EMBL/GenBank/DBJ databases.
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (probable).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF15053; AAK30684.1; -
CC Antigen.
KM NON TER 1 1
FT NON TER 1013 1013
SQ SEQUENCE 1013 AA; 110550 MW; 856E98912315D102 CRC64;

Query Match 6.9%; Score 95.5; DB 1; Length 1013;
Best Local Similarity 20.3%; Pred. No. 69;
Matches 59; Conservative 40; Mismatches 124; Indels 67; Gaps 9;

QY 31 ISKTAQITIKYGVAGPEQAAVAEVAQVAKENYL-----TVELV 70
DB 641 LKTNITQAITSNVLDGP--ATREVGKSLIQETNTVAGSSLEADQKAEIVKVGETTATH 698
QY 71 EENDVAMPNSAVSKGELDANAWQHPYV-EKDSQKGLNIVGVNTFVPLAGYSTKIX 129
DB 699 SDTSLSLRNKLLIMASAEKIVESTKNLPDRILMKGLVDGIYBKSGPELTAVSSGID 758
QY 130 TLNEIKDQATTIVPNDPSNLAALILKEKGLI-----KLKDNLT 170
DB 759 NSN-INDESEKALKKAKDAASEATLIDITONTLEGKQONTEHKKPRDDIYNKQAEVINA 817
QY 171 PSTTIDYENPKKLYTKE-----VDTSYAARAIDVDVLAANNVAGVGTASGVYF- 224
DB 818 VNPVLEALEKPAKPVSAERIVQETSSILNNISGLAEKVNPPA-----MLSSNGKPF 872
QY 225 -VEDKSDPYVNIIVARADNKSKAIDQPVKAYQDVEAEAKKQPKDQVI 273
DB 873 TLEKKKEESI-----KKYDELIVKAFGTGKSTEEQOSFKAMLI 910

RESULT 39
HAP1_HAEIN STANDARD; PRT; 1409 AA.
ID HAP1_HAEIN
AC P44596;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Adhesion and penetration protein precursor (EC 3.4.21.-).
GN HAP OR HI0248.
OS Haemophilus influenzae.
OC Bacteriae; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=9530630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shiley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Ulfersack T.R., Hanna M.C., Nguyen D.T., Saudek D.W., Brandon R.C.,
RA Fine L.D., Fritchman D.L., Fuhmann J.L., Geoghegan N.S.W.,

RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RL Rd."?
RL Science 269:496-512(1995).
CC -I- FUNCTION: PROBABLE PROTEASE. PROMOTES ADHERENCE AND INVASION BY
CC DIRECTLY BINDING TO A HOST CELL STRUCTURE (BY SIMILARITY).
CC -I- SUBCELLULAR LOCATION: Secreted (Potential).
CC -I- DOMAIN: THE SIGNAL PEPTIDE GUIDE THE PRECURSOR TO THE PERIPLASMIC
CC SPACE, AND THE CARBOXY-TERMINAL HELPER DOMAIN ASSOCIATES WITH THE
CC OUTER MEMBRANE TO FORM A PORE FOR EXCRETION OF THE PROTEASE
CC DOMAIN. THE HELPER DOMAIN IS THEN RELEASED BY AUTOPROTEOLYSIS (BY
CC SIMILARITY).
CC -I- SIMILARITY: Belongs to peptidase family S6.
CC -I- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; A STOP CODON HAD TO
CC BE SKIPPED IN POSITION 710 TO PRODUCE THIS ORF.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U32710; -; NOT_ANNOTATED_CDS.
CC -----
CC MEROPS; S06.006; -.
DR TIGR; HI0248; -.
DR InterPro; IPR006315; Autotransport.
DR InterPro; IPR005546; Autotransporter.
DR InterPro; IPR000710; Peptidase_S6.
DR InterPro; IPR004899; Peetractin.
DR Pfam; PF03797; Autotransporter; 1.
DR Pfam; PF02395; IGAI; 1.
DR Pfam; PF03212; Peetractin; 1.
DR PRINTS; PR00921; IGASERPTASE.
DR TIGRfams; TIGR01414; autotrans_bar1.1.
KW Hydrolyase, Serine protease, Transmembrane, Zymogen, Signal;
KW Complete proteome.
FT SIGNAL 1 25
FT CHAIN 26 ? POTENTIAL.
FT PROPEP ? ? ADHESION AND PENETRATION PROTEIN.
FT ACT SITE 250 250 HELPER PEPTIDE (POTENTIAL).
FT ACT SITE 250 250 BY SIMILARITY.
SQ SEQUENCE 1409 AA; 156797 MW; 63ABG937A84D16E CRC64;

Query Match 6.9%; Score 95.5; DB 1; Length 1409;
Best Local Similarity 23.9%; Pred. No. 1e+02;
Matches 53; Conservative 24; Mismatches 80; Indels 65; Gaps 9;

QY 64 NLTVLEVFNDVAMPNSAVSKGELDANAMQHKRYLEKD-----SQEGLNIVGVN----- 115
DB 852 NLTLN-----NSTVTLNSAVSAS--SNVAPRRRSLSLETPTTSAERFNTLVNGKLSGQ 905
QY 116 ---TFVYPLAGY-STKIKITNELKDGATTAVPNDPSNLAALILKEKGLIKLKDNTNLF 171
DB 906 GTQGTSSLTGYSKDKLKSNDABGDTLSV-----RNTGKEPTLRLDTLSLDNKPIS 961
QY 172 STTIDYENPKKLYKEVDTSVARAIDVDVLAANNVAGVGTASGVVEEDKSP 231
DB 962 DKLFTLEN-----DVAAGALRYKLVKNKGEFLNHPI----- 995
QY 232 YVNIIVARADNKSKAIDQPVKAYQD-----EVEAEAKKQ 267
DB 996 -----KEGELNDIVRAEQAQETLEAKQVQETAKKQ 1026

RESULT 40
VATC_MOUSE STANDARD; PRT; 382 AA.
ID VATC_MOUSE
AC Q921G3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Vacuolar ATP synthase subunit C (EC 3.6.3.14) (V-ATPase C subunit)
 DE (Vacuolar proton pump C subunit).
 GN ATP6V1C1 OR ATP6C OR V-ATC.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Howell M.L., Dean G.E.;
 RT "cDNA sequences for mouse vacuolar ATPase subunits";
 RL Submitted (Aug-1994) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: Subunit of the peripheral V1 complex of vacuolar ATPase.
 CC Subunit C is necessary for the assembly of the catalytic sector of
 CC the enzyme and is likely to have a specific function in its
 CC catalytic activity. V-ATPase is responsible for acidifying a
 CC variety of intracellular compartments in eukaryotic cells.

CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (in) = ADP + phosphate +
 CC H(+) (out).
 CC -1- SUBUNIT: V-ATPase is a heteromultimeric enzyme composed of a
 CC peripheral catalytic V1 complex (components A to H) attached to
 CC an integral membrane V0 proton pore complex (components: a, c, c',
 CC c'', and d).

CC -1- SIMILARITY: Belongs to the V-ATPase C subunit family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U13839; AAC83084.1; -

DR MGD: MG1:1913585; ATP6V1C1.

DR GO: GO:0008553; F:hydrogen-exporting ATPase activity, phospho. . .; IDA.

DR InterPro: IPR004907; V-ATPase_C.1.

DR Pfam: PF03223; V-ATPase_C.1.

DR Hydrolase; ATP synthetase; Hydrogen ion transport.

KW SEQUENCE 382 AA; 43860 MW; CDF51D6EC9439 CRC64;

Query Match 6.9%; Score 95; DB 1; Length 382;

Best Local Similarity 23.0%; Pred. No. 23; Indels 88; Gaps 16;

Matches 70; Conservative 46; Mismatches 100; Indels 88; Gaps 16;

QY 18 LAGCSNQSNEPAISKTAQTKVG--VMAGPEQVAEVAQVAKKYNLTVELVEFND 74
 Db 22 LHAATTKNNNLAVSSKFNIDLVKGTLDVVG-----LSDELAKLD- 62

QY 75 YAMPNSAVSKGELDANAMQKFPYLEKDSQKGLNNLVIVG-----NTFYIPLAGYSTK 127
 Db 63 -AFVEGVVK-----VAQYMDVLEDSKQVQENLNASGVDLVTYITRFQMDVAKYPIK 115

QY 128 --IKTLNELKDGATIAVPNDPSNLARALLLEKQGLIKLKQNTNLF--TLLDIYENPKL 184
 Db 116 QSLKNTSEITIAKGVQIDNDLKSRSAYNNL--KGNLQNLERKNAGSLTSLAE----- 168

QY 185 VIKENDTSVAARAIDVDLAVV-----NNVAGQVGLTAS-----ENGVFVEDKDSPTVN 234
 Db 169 IYKQDFVLDSEYL--VTLVVVVKPLNNDWIKQYETLAEMVPPRSSNVLSDDQSYLCN 226

QY 235 IIVARADNDSKAIDF-----VYAYQTDVEVEAA-----KKQFKQGV 272
 Db 227 VTLFPR-----KAVDDPRKARENKFIVADFQYNEEMKADKEENTELSTDKKQF--GP 278

QY 273 IKGM 276
 Db 279 LVRW 282

Search completed: June 16, 2004, 11:07:10
 Job time : 21 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 16, 2004, 11:04:11 ; Search time 46 Seconds
(without alignments)
1893.109 Million cell updates/sec

Title: US-10-018-672-2
Perfect score: 1383
Sequence: 1 MNPFGKINGICAMASIALAG.....TDEYEAERKKQPKDGVIKGM 276

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 segs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL 25: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phage: *
10: sp_plant: *
11: sp_ricent: *
12: sp_virus: *
13: sp_vertebrate: *
14: sp_unclassified: *
15: sp_virus: *
16: sp_bacteriophage: *
17: sp_archaeal: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	958.5	69.3	270	16	Q7VM93	Q7VM93 haemophilus
2	786.5	56.9	271	16	Q83MC6	Q83MC6 shigella fl
3	770	55.7	270	16	Q8CWC8	Q8CWC8 escherichia
4	752.5	54.4	269	16	Q8DFC1	Q8DFC1 vibrio vuln
5	739.5	53.5	269	16	Q87R33	Q87R33 vibrio para
6	685.5	49.6	272	16	Q8CVC2	Q8CVC2 escherichia
7	571.5	41.3	267	16	Q83F42	Q83F42 coxiella bu
8	571	41.3	259	2	Q69441	Q69441 legionella
9	563.5	40.7	266	16	Q8Y0X1	Q8Y0X1 ralsionia s
10	560.5	40.5	278	16	Q8YD39	Q8YD39 bruceella me
11	560.5	40.5	278	16	Q8FV86	Q8FV86 bruceella su
12	554	40.1	262	16	Q7W317	Q7W317 bordetella
13	554	40.0	262	16	Q7W317	Q7W317 bordetella
14	551	39.8	262	16	Q7VSM1	Q7VSM1 bordetella
15	527.5	38.1	265	16	Q7WV1	Q7WV1 bordetella
16	527.5	38.1	265	16	Q7W4E3	Q7W4E3 bordetella

17	526.5	38.1	265	16	Q7VV70	Q7VV70 bordetella
18	524.5	37.9	258	16	Q921X5	Q921X5 rhizobium m
19	517.5	37.4	259	16	Q8XU08	Q8XU08 ralsionia s
20	517	37.4	261	16	Q8RFN4	Q8RFN4 fusobacteri
21	517	37.4	261	16	Q8RFD1	Q8RFD1 rhizobium l
22	515.5	37.3	259	16	Q8U7G0	Q8U7G0 agrobacteri
23	513	37.1	307	16	Q89E28	Q89E28 bradyrhizob
24	509.5	36.8	286	16	Q9CIN7	Q9CIN7 lactococcus
25	507	36.7	259	16	Q8FXS4	Q8FXS4 bruceella su
26	507	36.7	268	16	Q8YEC4	Q8YEC4 bruceella me
27	506.5	36.6	261	16	Q88CL5	Q88CL5 pseudomonas
28	502.5	36.3	260	16	Q9HT68	Q9HT68 pseudomonas
29	501	36.2	269	16	Q8PGF0	Q8PGF0 xanthomonas
30	497.5	36.0	259	16	Q8XHS5	Q8XHS5 pseudomonas
31	497.5	36.0	266	16	Q8P4S9	Q8P4S9 xanthomonas
32	488	35.3	256	16	Q88RL7	Q88RL7 pseudomonas
33	487	35.2	257	16	Q87UN6	Q87UN6 pseudomonas
34	485	35.1	268	16	Q7V195	Q7V195 helicobacte
35	481	34.8	277	16	Q832Y8	Q832Y8 enterococcu
36	477.5	34.5	257	16	Q9PEP7	Q9PEP7 campylobact
37	469.5	33.9	273	16	Q8ELA4	Q8ELA4 oceanobacti
38	468	33.8	256	16	Q9PEP8	Q9PEP8 campylobact
39	467.5	33.8	271	16	Q26084	Q26084 helicobacte
40	464.5	33.6	286	16	Q9CIN6	Q9CIN6 lactococcus
41	463.5	33.5	273	16	Q92E25	Q92E25 listeria in
42	462.5	33.4	271	16	Q92U45	Q92U45 helicobacte
43	462.5	33.4	273	16	Q8YA74	Q8YA74 listeria mo
44	460	33.3	286	16	Q9CIN8	Q9CIN8 lactococcus
45	457.5	33.1	260	16	Q87UT0	Q87UT0 pseudomonas

ALIGNMENTS

RESULT 1

ID	Q7VM93	PRELIMINARY	FRT	270 AA.
AC	Q7VM93;			
DT	01-OCT-2003 (TREMBLrel. 25, Created)			
DT	01-OCT-2003 (TREMBLrel. 25, Last sequence update)			
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)			
DE	28kDa outer membrane lipoprotein.			
GN	HLPA OR HD1100.			
OS	Haemophilus ducreyi.			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;			
OC	Pasteurellaceae; Haemophilus.			
OX	NCBI_TaxID=730;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=3500HP / ATCC 700724;			
RA	Munson R.S., Jr., Ray W.C., Manaitas G., Sabo P., Mungur R.,			
RA	Johnson L., Nguyen D., Wang J., Forest C., Hood L.,			
RT	"The complete genome sequence of Haemophilus ducreyi."			
RL	Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.			
DR	EMBL, AF017153; AAP95966.1; -			
KM	lipoprotein; Complete proteome.			
SQ	SEQUENCE 270 AA; 29535 MW; FC077F10B13304AD CRC64;			

QY	Query Match	69.3%; Score 958.5; DB 16; Length 270;
QY	Best Local Similarity	71.5%; Pred. No. 1.4e-53;
DB	Matches 188; Conservative 34; Mismatches 36; Indels 5; Gaps 2;	
QY	14	SCIALAGSNGSNEPRAISKTAQITKGVAGQQAFAEYAGVAKKYLTVELVEFN 73
DB	13	STLALTGCKPANNQ---GDSASTIKIKGVMSGPHSVAEKRAATKINLDEVEFYLFN 68
QY	74	DYAMNSAVSKGELDANAMQKPYLEKDSQKGLNLTIVAGTFYYPPLAGYSTIKITLNE 133
DB	69	DYALFNVAVSKGDDIDNMQKPYLDKDSQKGLNLTIVAGTFYYPPLAGYSTIKITLNE 128
QY	134	LKDQATIVPNDPSNLALILILEKQGIKIKDNTNLFSTLIDIVENKTKIVIEVDTSV 193
DB	129	LKSEAVVAVPNDPSNHARALMLLEKQGITLINDTNLFATKLDIVDNKTKILKEVDTSV 188

QY 194 ARAIDVDVLA VNNNTAGVGLTASENGVEVEDKDS PYNNITVARADNKSALDQDFYK 253
 DB 189 AAKLDDVDLA VNNNTASQAGLSVNS-VFVEDKDS PYNNITVARADNKSALDQDFYK 247
 QY 254 AYQTEVEAEAKQFKDGVTKGW 276
 DB 248 AYQTEVEAEAKQFKDGVTKGW 270

RESULT 2

Q83MC6 PRELIMINARY; PRT; 271 AA.
 AC Q83MC6;
 DT 01-JUN-2003 (Tremblrel. 24, Created)
 DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Putative lipoprotein.
 GN YAC OR SF0188 OR S0190.
 OS Shigella flexneri.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Shigella.
 CX NCBI_Taxid=623;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=301 / Serotype 2a;
 RX MEDLINE=2272406; PubMed=12384590;
 RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Ju W., Wang J., Liu H.,
 RA Yang J., Yang F., Zhang J., Zhang G., Wu H., Qu D., Dong J.,
 RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
 RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
 RA Yu J.,
 RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
 RT through comparison with genomes of Escherichia coli K12 and O157.";
 RL Nucleic Acids Res. 30:4432-4441(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
 RX MEDLINE=22590274; PubMed=12704152;
 RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
 RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
 RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
 RA Schwartz D.C., Blattner F.R.;
 RT "Complete genome sequence and comparative genomics of Shigella
 RT flexneri serotype 2a strain 2457T.";
 RL Infect. Immun. 71:2775-2786(2003).
 DR EMBL; AB015055; AA041850.1; -;
 DR EMBL; AB016978; AA015730.1; -;
 DR InterPro; IPR004872; Lipoprotein_9.
 DR InterPro; IPR000437; Prok_lipoprot_S.
 DR InterPro; IPR004478; YaeC_
 DR Pfam; PF03180; Lipoprotein_9.1.
 DR TIGRFAMs; TIGR00363; TIGR00363.1.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 DR Lipoprotein; Complete proteome.
 SQ SEQUENCE 271 AA; 29417 MW; C3D7D01FB5CD2BF6 CRC64;

Query Match 56.9%; Score 786.5; DB 16; Length 271;
 Best Local Similarity 55.4%; Pred. No. 1,4e+42;

Matches 133; Conservative 45; Mismatches 67; Indels 11; Gaps 3;

QY 3 FKGINGICALASGIALAGCSNQNNEPAISKTAQTIKGVMAQGEQAVAEVAGVAXEK 62
 DB 5 FKTFAAVGALIGSLALVGCQDEKDP-----NHIKGVIVGAEQVAVAGVAXEK 56
 QY 63 YNLTVELVFNNDYAMPNSAVSKGELDANMCHPYLEKDSQKGLNNLVYGVTFYPLA 122
 DB 57 YGIDVELVTFNDYVLENEALSKGDDIDANAFQHKPYLDQOLKDRGY-KLVAAGNTFVYPIA 115
 QY 123 GYSTIKITNELKDGATIAVPNDPSNLARALILEKQGLIKLKDNTNLFSTLDIVENPK 182
 DB 116 GYSKIKISLDELQDSQVAVPNDPTNLGSLILLIKVGLIKLKDQVGLLPTVLDAVENPK 175

QY 163 KLVIKEVDTSVAARAID--VDLAVNNNTAGVGLTASENGVEVEDKDS PYNNITVARA 240
 DB 176 NKLIVELAPQLPNSLDAQIALAVINTTASQIGLTPAKGIVFEDKDS PYNNITVIRE 235
 QY 241 DNKDSKAIQDFYKAYQTEVEAEAKQFKDGVTKGW 276
 DB 236 DNKDAENYKKEFVQAYQSDVEAEANKVFNGAVTKGW 271

RESULT 3

Q8CWC8 PRELIMINARY; PRT; 270 AA.
 AC Q8CWC8;
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE D-mechonine-binding lipoprotein mecQ precursor.
 GN YAC OR C0238.
 OS Escherichia coli O6.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 CX NCBI_Taxid=217992;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O6:H1 / CFT073 / ATCC 700928;
 RX MEDLINE=22388234; PubMed=12471157;
 RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
 RA Raeko D., Buckles E.L., Liou S.-R., Boutin A., Hackert J., Stroud D.,
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
 RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
 RT "Extensive mosaic structure revealed by the complete genome sequence
 RT of uropathogenic Escherichia coli.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
 DR EMBL; AB016755; AA078730.1; -;
 DR InterPro; IPR004872; Lipoprotein_9.
 DR InterPro; IPR000437; Prok_lipoprot_S.
 DR InterPro; IPR004478; YaeC_
 DR Pfam; PF03180; Lipoprotein_9.1.
 DR TIGRFAMs; TIGR00363; TIGR00363.1.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 DR Lipoprotein; Complete proteome.
 SQ SEQUENCE 270 AA; 29318 MW; 2EC4133C671EALBC CRC64;

Query Match 55.7%; Score 770; DB 16; Length 270;
 Best Local Similarity 55.1%; Pred. No. 1,6e+41;

Matches 152; Conservative 45; Mismatches 67; Indels 12; Gaps 4;

QY 3 FKGINGICALASGIALAGCSNQNNEPAISKTAQTIKGVMAQGEQAVAEVAGVAXEK 62
 DB 5 FKTFAAVGALIGSLALVGCQDEKDP-----NHIKGVIVGAEQVAVAGVAXEK 56
 QY 63 YNLTVELVFNNDYAMPNSAVSKGELDANMCHPYLEKDSQKGLNNLVYGVTFYPLA 122
 DB 57 YGIDVELVTFNDYVLENEALSKGDDIDANAFQHKPYLDQOLKDRGY-KLVAAGNTFVYPIA 115
 QY 123 GYSTIKITNELKDGATIAVPNDPSNLARALILEKQGLIKLKDNTNLFSTLDIVENPK 182
 DB 116 GYSKIKISLDELQDSQVAVPNDPTNLGSLILLIKVGLIKLKDQVGLLPTVLDAVENPK 174
 QY 183 KLVIKEVDTSVAARAID--VDLAVNNNTAGVGLTASENGVEVEDKDS PYNNITVARA 240
 DB 175 NKLIVELAPQLPNSLDAQIALAVINTTASQIGLTPAKDGI FVDEKDS PYNNITVIRE 234
 QY 241 DNKDSKAIQDFYKAYQTEVEAEAKQFKDGVTKGW 276
 DB 235 DNKDAENYKKEFVQAYQSDVEAEANKVFNGAVTKGW 270

RESULT 4

Q8DFC1 PRELIMINARY; PRT; 269 AA.
 AC Q8DFC1;
 DT 01-MAR-2003 (Tremblrel. 23, Created)

DT 01-MAR-2003 (Tremblrel. 23, last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, last annotation update)
 DE ABC-type metal ion transport system, periplasmic component.
 GN Vli0294.
 OS *Vibrio vulnificus*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; *Vibrio*.
 NCBI_Taxid=672;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CMC86;
 RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
 RA Choy H.E.;
 RT "Complete genome sequence of *Vibrio vulnificus* CMC86";
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB016797; AAC08827.1; -;
 DR InterPro; IPR004872; Lipoprotein_9.
 DR InterPro; IPR004437; Prok_Lipoprot_9.
 DR InterPro; IPR004478; Yaec.
 DR Pfam; PF03180; Lipoprotein_9; 1.
 DR TIGRPFAMs; TIGR00363; TIGR00363; 1.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 DR Complete proteome.
 KW SEQUENCE 269 AA; 29234 MW; 774C7112BC3F1E15 CRC64;
 SQ
 Query Match 54.4%; Score 752.5; DB 16; Length 269;
 Best Local Similarity 54.5%; Pred. No. 2e-40;
 Matches 146; Conservative 44; Mismatches 69; Indels 9; Gaps 2;
 QY 9 ICALASGIALAGCSNNSPPAISKTAQITIKYGVAGPEQAVAVAGVAAEKXNLTVE 68
 DB 11 IATAASLALVLAGCGDK-----EVDVNVKVGVIAGAEQAVAAKVAEKXNLTVE 62
 QY 69 LVEFNDYAMPNSAVSKGLDANAMQHPYLEKDSQKGLNNLVIGNTFYVPLAGYSTKI 128
 DB 63 LVTFDYVTPNALDDGSIDINAFQHKPYLDQVDRDY-KLTIGNTFYVPIAGYSTKY 121
 QY 129 KTLNELKDGATIAVPNDPSNLARALILKQGLIKLKDNTNLFSTLTDIENPKKLVKE 188
 DB 122 KSVDELQGARIAVPNDPNTLGRSLTLLEQGLITLSDVGLATFVDIVGNPKNLTVE 181
 QY 189 VDTSVARAIDVDVLA VNNNNYAGVGLTASENGVEFDKSPYVNIIVARADNDSKAI 248
 DB 182 LPAQQLPRSLDDVTLTIINTYASIDLSPEKDGVEFDKSPYVNIIVAREVNVANQV 241
 QY 249 QDFKAYQTEVEVEAEAKKQFQDGVTKGW 276
 DB 242 QNFVKAQTEVEVEAEAKKIFKGGVVGW 269
 RESULT 5
 QSTRS3 PRELIMINARY; PRT; 269 AA.
 ID Q87RS3;
 AC Q87RS3;
 DT 01-JUN-2003 (Tremblrel. 24, Created)
 DT 01-JUN-2003 (Tremblrel. 24, last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, last annotation update)
 DE Lipoprotein Yaec.
 GN Vli0704.
 OS *Vibrio parahaemolyticus*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; *Vibrio*.
 NCBI_Taxid=670;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RIND 2210653 / Serotype O3:K6;
 RA MEDLINE=22508454; PubMed=12620739;
 RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
 RA Iijima Y., Naito M., Nakano M., Yamashita A., Kubota Y., Kilmura S.,
 RA Yasunaga T., Honda T., Shingawa H., Hattori M., Iida T.;
 RT "Genome sequence of *Vibrio parahaemolyticus*: a pathogenic mechanism
 distinct from that of *V. cholerae*.";
 RL Lancet 361:743-749 (2003).

DR EMBL; AP005075; BACS8967.1; -;
 DR InterPro; IPR004872; Lipoprotein_9.
 DR InterPro; IPR004437; Prok_Lipoprot_9.
 DR Pfam; PF03180; Lipoprotein_9; 1.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 DR Complete proteome.
 KW SEQUENCE 269 AA; 29070 MW; 5D458EA921862ED8 CRC64;
 SQ
 Query Match 53.5%; Score 739.5; DB 16; Length 269;
 Best Local Similarity 52.6%; Pred. No. 1.4e-39;
 Matches 141; Conservative 52; Mismatches 66; Indels 9; Gaps 2;
 QY 9 ICALASGIALAGCSNNSPPAISKTAQITIKYGVAGPEQAVAVAGVAAEKXNLTVE 68
 DB 11 VATAASLALVLAGCGDK-----EVDVNVKVGVIAGAEQAVAAKVAEKXNLTVE 62
 QY 69 LVEFNDYAMPNSAVSKGLDANAMQHPYLEKDSQKGLNNLVIGNTFYVPLAGYSTKI 128
 DB 63 LVTFDYVTPNALDDGSVDANAFQHKPYLDQVDRDY-KLAIAGNSFYVPIAGYSTKY 121
 QY 129 KTLNELKDGATIAVPNDPSNLARALILKQGLIKLKDNTNLFSTLTDIENPKKLVKE 188
 DB 122 KSVDELQGARIAVPNDPNTLGRSLTLLEQGLITLSDVGLATFVDIVGNPKNLTVE 181
 QY 189 VDTSVARAIDVDVLA VNNNNYAGVGLTASENGVEFDKSPYVNIIVARADNDSKAI 248
 DB 182 LEAQQLPRSLDDVTLTIINTYASIDLSPEKDGVEFDKSPYVNIIVAREVNVANQV 241
 QY 249 QDFKAYQTEVEVEAEAKKQFQDGVTKGW 276
 DB 242 QNFVKAQTEVEVEAEAKKIFKGGVVGW 269
 RESULT 6
 Q8CVK2 PRELIMINARY; PRT; 272 AA.
 ID Q8CVK2;
 AC Q8CVK2;
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, last annotation update)
 DE Lipoprotein-28 precursor.
 GN NIBA OR C4583.
 OS *Escherichia coli* O6.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; *Escherichia*.
 NCBI_Taxid=217992;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O6:H1 / ATCC 700928;
 RA MEDLINE=22386234; PubMed=12471157;
 RA Welch R.A., Burland V., Plunkett G. III, Redford P., Resch P.,
 RA Raeko D., Buckles E.L., Liu S.-R., Boutin A., Hackett J., Stroud D.,
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
 RA Mobley H.L.T., Domenech M.S., Blattner F.R.;
 RT "Extensive mosaic structure revealed by the complete genome sequence
 of uropathogenic *Escherichia coli* O6".
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
 DR EMBL; AE016769; AAN83017.1; -;
 DR InterPro; IPR004872; Lipoprotein_9.
 DR InterPro; IPR004437; Prok_Lipoprot_9.
 DR InterPro; IPR004478; Yaec.
 DR Pfam; PF03180; Lipoprotein_9; 1.
 DR TIGRPFAMs; TIGR00363; TIGR00363; 1.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 DR Complete proteome.
 KW SEQUENCE 272 AA; 29414 MW; B0CFAD9022B3D2F4 CRC64;
 SQ
 Query Match 49.6%; Score 685.5; DB 16; Length 272;
 Best Local Similarity 49.8%; Pred. No. 3.9e-36;
 Matches 135; Conservative 54; Mismatches 71; Indels 11; Gaps 3;
 QY 8 GICALASGIALAGCSNNSPPAISKTAQITIKYGVAGPEQAVAVAGVAAEKXNLTVE 67

[1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GM1000;
 RX MEDLINE=21681879; PubMed=11823852;
 RA Salanoubat M., Genin S., Artiguenave F., Guzy J., Mangenot S.,
 RA Arlat M., Billault A., Broctier P., Camus J.C., Catolico L.,
 RA Chanderler C., Choise N., Claudel-Renard C., Cunac S., Demange N.,
 RA Gaspin C., Lave M., Moisan A., Robert C., Saurin W., Schlex T.,
 RA Siquier P., Thebaud P., Whalen M., Wincker P., Levy M.,
 RA Weisenbach J., Boucher C.A.;
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum";
 RL Nature 415:497-502(2002).
 DR EMBL: AL646061; CAD14624.1; -
 DR InterPro: IPR004872; Lipoprotein_9.
 DR InterPro: IPR006311; Tat.
 DR InterPro: IPR004478; YaeC.
 DR Pfam: PF03180; Lipoprotein_9; 1.
 DR TIGRFAMs: TIGR01409; Tat signal_seq; 1.
 DR TIGRFAMs: TIGR00363; TIGR00363; 1.
 KM Complete proteome.
 SQ SEQUENCE 266 AA; 28515 MW; B718F52C9CF9B236 CRC64;

Query Match 40.7%; Score 563.5; DB 16; Length 266;
 Best Local Similarity 43.2%; Pred. No. 2.3e-28;
 Matches 115; Conservative 47; Mismatches 87; Indels 17; Gaps 3;

QY 11 ALASGIALAGCSNQSNEPAISKTAQTIKVGWAGPEQAVAEVAGQVAKKYNLTVELV 70
 DB 18 ALAAGALG-----CGKPIKIGVLAGPHEIMEAVKVA-EGDGLKQIV 61
 QY 71 EFNDYAMPNSAVSKGELDANAMQHKPYLEKDSQKGLNLTIVGNTFYPLAGYSTIKIT 130
 DB 62 EFNDYIOPRAALAGGLDANSYOHQPYLDDQATATKCY-KFVSVGQITTFPMGVSKKKS 120
 QY 131 LNELKDGATIVPNDPSNLARALILLEKQGLIKLKDNTNLFSTLTIENPKLVKEVD 190
 DB 121 LNDLKDGARFGFNDPTNGRALLLLQAGVILKKNAGFKASPRDVAENPKLKFVELD 180
 QY 191 TSVARATIDVDLAVNNNNYAGVGLTASENGVFVEDKDSPTVNTIIVAPADNCKSKAIQ 250
 DB 181 AAQLPRLSDLDLAAVNGVYAKAGLDPRDGLAIESPKPYANVIAVADKQDPYAK 240
 QY 251 FVKAYQTDVEAEAKKQFKDGVYIKGW 276
 DB 241 LYKAVHSDAVKAFVTKYKDVAVIYAM 266

RESULT 10

Q8YD39 PRELIMINARY; PRT; 278 AA.
 AC Q8YD39;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE ABC transporter substrate binding protein.
 GN BMEI10338.
 OS Brucella melitensis.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 CC Brucellaceae; Brucella.
 CX NCB1_TaxID=29459;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=16M / ATCC 23456 / Biotype 1;
 RX MEDLINE=20020109; PubMed=11756688;
 RA Delvecchio V.G., Kapetral V., Redkar R.J., Patra G., Wujer C., Los T.,
 RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
 RA Jablonki L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,
 RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Lelesson J.-D.,
 RA Haeflorn R., Kyplides N., Overbeek R.;
 RT "The genome sequence of the facultative intracellular pathogen
 RT Brucella melitensis";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
 DR EMBL: AB009671; AAL53580.1; -

DR PIR: A13551; A13551.
 DR InterPro: IPR004872; Lipoprotein_9.
 DR InterPro: IPR006311; Tat.
 DR InterPro: IPR004478; YaeC.
 DR Pfam: PF03180; Lipoprotein_9; 1.
 DR TIGRFAMs: TIGR01409; Tat signal_seq; 1.
 DR TIGRFAMs: TIGR00363; TIGR00363; 1.
 KM Complete proteome.
 SQ SEQUENCE 278 AA; 30606 MW; C0126189CC59B362 CRC64;

Query Match 40.5%; Score 560.5; DB 16; Length 278;
 Best Local Similarity 44.2%; Pred. No. 3.8e-28;
 Matches 118; Conservative 50; Mismatches 84; Indels 15; Gaps 5;

QY 11 ALASGIALAGCSNQSNEPAISKTAQTIKVGWAGPEQAVAEVAGQVAKKYNLTVELV 70
 DB 25 ALTVFA-----SAPSHAEKTKIKVINGGEDEDVWKAVERGK-KHGINTERI 72
 QY 71 EFNDYAMPNSAVSKGELDANAMQHKPYLEKDSQKGLNLTIVGNTFYPLAGYSTIKIT 130
 DB 73 EFNDYQNEALEKEIDANAFQHKPYLDEQIKQHG-KISVAGTYAVPFGIYSRKVK 131
 QY 131 LNELKDGATIVPNDPSNLARALILLEKQGLIKLKDNTNLFSTLTIENPKLVKEVD 190
 DB 132 LNELKDGATVGVFNDPTNGRALLVLEQGLIKLKPDAGIATPDIENPKLKEIYELD 191
 QY 191 TSVARATIDVDLAVNNNNYAGVGLTASENGVFVEDKDSPTVNTIIVAPADNCKSKAIQ 249
 DB 192 AGVGRSITDDLDALIVNNDAKAGL-KKEDATGESKKNPNYNTAVRTDLDQPMVK 250
 QY 250 DFVKAYQTDVEAEAKKQFKDGVYIKGW 276
 DB 251 ALVASFQNDVAKELERAYKGTGIPAW 277

RESULT 11

Q8FV86 PRELIMINARY; PRT; 278 AA.
 AC Q8FV86;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Lipoprotein, YaeC family.
 GN BR00960.
 OS Brucella suis.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 CC Brucellaceae; Brucella.
 CX NCB1_TaxID=29461;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1330 / Biovar 1;
 RX MEDLINE=22247741; PubMed=12271122;
 RA Paulsen I.T., Seshadri R., Nelson K.E., Bisen J.A., Heidelberg J.F.,
 RA Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beanan M.J.,
 RA Daugherty S.C., Debey R.T., Durkin A.S., Kolonay J.F., Madupu R.,
 RA Nelson W.C., Ayodele J.B., Krahl M., Shetty J., Malek J., Van Aken S.E.,
 RA Riedmiller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,
 RA Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;
 RT "The Brucella suis genome reveals fundamental similarities between
 RT animal and plant pathogens and symbionts";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
 DR EMBL: AE014589; AAN34131.1; -
 DR PIR: A13551; A13551.
 DR TIGR: BR00960; -
 DR InterPro: IPR004872; Lipoprotein_9.
 DR InterPro: IPR006311; Tat.
 DR InterPro: IPR004478; YaeC.
 DR Pfam: PF03180; Lipoprotein_9; 1.
 DR TIGRFAMs: TIGR01409; Tat signal_seq; 1.
 DR TIGRFAMs: TIGR00363; TIGR00363; 1.
 KM Complete proteome.
 SQ SEQUENCE 278 AA; 30606 MW; C0126189CC59B362 CRC64;

Query Match 40.5%; Score 560.5; DB 16; Length 278;
 Best Local Similarity 44.2%; Pred. No. 3.8e-26;
 Matches 118; Conservative 50; Mismatches 84; Indels 15; Gaps 5;

QY 11 ALASGIALAGCSNOSNEPAISKTAQOTIKYGVNAGBQAVAEVAGVAKXNLTVEV 70
 DB 25 ALTVGFA-----SAPSHADKTIKIGVIGGDEDEVMKVAEKG-KGLNIERI 72
 QY 71 EENDYAMNSAVSKGELDANAMOHKPYLEKDSQKGLNINVIYVNTVYLAQYRK 130
 DB 73 TENDYQPEALEREKIDANAFQHKPYDEQIKOHGY-KISVAGYIVMPTIGYSRKVK 131
 QY 131 INELKDGATIAVNDPSNLARALILEKQGLIKLKDNTNFSTLIDIVENPKLV 190
 DB 132 LEELKDGATVGVNDPTNEGRALRVLEBQGLIKLKPAIGATIPIDIVENPKL 191
 QY 191 TSVAARATDDVDLAVNNNVAQGVGTASENGVEVEKD-SPYVNIIVAADKDSK 249
 DB 192 AGVGRSISDDLDALAVNDMAKAGL-KKEDALIKESKNNPYNFIATVDLDDQ 250
 QY 250 DFVAKYQTEVEAEAKKQFQKGV 276
 DB 251 ALVASFQNDVAVKAELERAYKGTGIPAW 277

RESULT 12
 ID Q7W317 PRELIMINARY; PRT; 262 AA.
 AC Q7W317;
 DT 01-OCT-2003 (TRENBLrel. 25, Created)
 DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Putative exported protein.
 GN BPA046
 OS Bordetella parapertussis.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 OX NCBI_TaxId=519;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=12822 / ATCC BAA-588;
 RX MEDLINE=22827954; PubMed=12910271;
 RA Parkhill U., Sebailia M., Preston A., Murphy L.D., Thomson N.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cerdeno-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Fellwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
 RA Leather S., Moutie S., Norbertczak H., O'Neill S., Ormond D., Price C.,
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
 RA Sharp S., Simmonds M., Skelton U., Squares R., Squares S., Stevens K.,
 RA Unwin L., Whitehead S., Bartell B.G., Maskell D.J.,
 RT "Comparative analysis of the genome sequences of Bordetella pertussis,
 RT Bordetella parapertussis and Bordetella bronchiseptica.",
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL, BX640435; CAB39329.1; -
 KM Complete proteome.
 SQ SEQUENCE 262 AA; DAC18652307E8B2A CRC64;

Query Match 40.1%; Score 554; DB 16; Length 262;
 Best Local Similarity 42.5%; Pred. No. 9.2e-28;
 Matches 116; Conservative 55; Mismatches 86; Indels 16; Gaps 4;

QY 1 MNFKINGICATAGSIALAGCSNOSNEPAISKTAQOTIKYGVNAGBQAVAEVAGVAK 60
 DB 3 MNFVR--SALLASAPFLAGAAQ-----AEKLVGATQVPHAEILEVVKAL 47
 QY 61 EKYNLTVEVFNQYAMPNSAVSKGELDANAMOHKPYLEKDSQKGLNINVIYVNT 120
 DB 48 AKGVYELDIKVFYVQPNQLADKQDANFQHPQYLDFTFNKDR-KTNIVSVGLVHVEP 106
 QY 121 LAGYSTRKTLNELKDGATIAVNDPSNLARALILEKQGLIKLKDNTNFSTLIDIVEN 180

DB 107 FGVYKKIKSLAEKDGATIAIPNDPSNGRALLLLQKQGLIKLKQDPSNIVATPIDIAEN 166
 QY 181 PKLVIEVDTSVAARATDDVDLAVNNNVAQGVGTASENGVEVEKDSFYVNIIVAA 240
 DB 167 PKKIFRELEAAMLRSLDLDLALINTVYALEAGVFTPRDLFEGADSPYANLVAA 226
 QY 241 DNKDSKATODFYKAYQTEVEAEAKKQFQKGV 273
 DB 227 DNKDAVAVKLVNALSSEAVRKFIIEKYKAVV 259

RESULT 13
 ID Q7WEW2 PRELIMINARY; PRT; 262 AA.
 AC Q7WEW2;
 DT 01-OCT-2003 (TRENBLrel. 25, Created)
 DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Putative exported protein.
 GN BPA519
 OS Bordetella bronchiseptica (Alcaligenes bronchiseptica).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 OX NCBI_TaxId=518;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RB50 / ATCC BAA-588;
 RX MEDLINE=22827954; PubMed=12910271;
 RA Parkhill U., Sebailia M., Preston A., Murphy L.D., Thomson N.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cerdeno-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Fellwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
 RA Leather S., Moutie S., Norbertczak H., O'Neill S., Ormond D., Price C.,
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
 RA Sharp S., Simmonds M., Skelton U., Squares R., Squares S., Stevens K.,
 RA Unwin L., Whitehead S., Bartell B.G., Maskell D.J.,
 RT "Comparative analysis of the genome sequences of Bordetella pertussis,
 RT Bordetella parapertussis and Bordetella bronchiseptica.",
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL, BX640450; CAB34882.1; -
 KM Complete proteome.
 SQ SEQUENCE 262 AA; 0DC80B1739FF4663 CRC64;

Query Match 40.0%; Score 553; DB 16; Length 262;
 Best Local Similarity 43.1%; Pred. No. 1.1e-27;
 Matches 116; Conservative 53; Mismatches 85; Indels 18; Gaps 5;

QY 1 MNFKINGICATAGSIALAGCSNOSNEPAISKTAQOTIKYGVNAGBQAVAEVAGVAK 59
 DB 3 MNFVR--SALLASAPFLAGAAQ-----AEKLVGATQVPHAEILEVVKPALA 48
 QY 60 KEKYNLTVEVFNQYAMPNSAVSKGELDANAMOHKPYLEKDSQKGLNINVIYVNT 119
 DB 49 KEGVYELDIKVFYVQPNQLADKQDANFQHPQYLDFTFNKDR-KTNIVSVGLVHVE 105
 QY 120 PLAGYSTRKTLNELKDGATIAVNDPSNLARALILEKQGLIKLKDNTNFSTLIDIVE 179
 DB 106 PFGYSTRKTLNELKDGATIAIPNDPSNGRALLLLQKQGLIKLKQDPSNIVATPIDIAE 165
 QY 180 NPKLVIEVDTSVAARATDDVDLAVNNNVAQGVGTASENGVEVEKDSFYVNIIVAA 239
 DB 166 NPKKIFRELEAAMLRSLDLDLALINTVYALEAGVFTPRDLFEGADSPYANLVAA 225
 QY 240 ADNKDSKATODFYKAYQTEVEAEAKKQFQKGV 273
 DB 226 PDNKDAVAVKLVNALSSEAVRKFIIEKYKAVV 259

RESULT 14
 ID Q7WSM1 PRELIMINARY; PRT; 262 AA.

```

AC Q7VSM1;
DT 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DE Putative exported protein.
GN BP0385.
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
NCBI_Taxid=520;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tohama 1 / ATCC BAA-589 / NCTC 13251;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebailia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Cherevach I.,
RA Achman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett U.,
RA Felwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
RA Leather S., Moulé S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica."
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640412; CAE44716.1; -.
KW Complete proteome.
SQ SEQUENCE 262 AA; 28524 MW; E0A56B1A59FE4C9 CRC64;

Query Match 39.8%; Score 551; DB 16; Length 262;
Best Local Similarity 43.1%; Pred. No. 1.4e-27;
Matches 118; Conservative 52; Mismatches 86; Indels 18; Gaps 5;

QY 1 MNFGKINGICALASGIALACGNSQNSNEPAISKTAQTIKVGVMGPEQAVNA-GOVA 59
DB 3 MNFVR---SALIASAFLAGAAQ-----AEKLVGATQVPHAEIIEVKKPALA 48
QY 60 KKKNTLTVLVEFNDYAMPNSAVSKGELDANAMQHKPYLEKDSQEKGLNLTIVGNTFY 119
DB 49 KGVVELDIKV--FTDVOQNLQDLADKQDANFQHPYIDTFNKOR-KNTLVSVGLVHYE 105
QY 120 PLAGVSTKIKTINELKDGATIVPNDPSNLARALITLKEQGLIKDKDNTLSTLDDIVE 179
DB 106 PFGGSIKSKIKSLAELKDGATITAIIPNDPSNGRALLKQGLKLDKDPNIVATPIDIAE 165
QY 180 NPKKLVKEVDTSVAAARAIDVDLAVVNNVYAGQVLTASENGVEVEDSPYVNIIVAR 239
DB 166 NPKKLFRELEAAMPFRSFDLDLALINTNVYALBGLVETRRALFLBAGDSYANLVAR 225
QY 240 ANDKSKAQDQFVKAYQTEVEAEAKKQFQDGYI 273
DB 226 PDKDAVPAVKLVNLSAEVAKETIEKKYGAIV 259

RESULT 15
Q7WFW1
ID Q7WFW1 PRELIMINARY; PRT; 265 AA.
AC Q7WFW1;
DT 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DE Putative exported protein.
GN B04169.
OS Bordetella bronchiseptica (Alcaligenes bronchiseptica).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
NCBI_Taxid=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271;

```

```

RA Parkhill J., Sebailia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Cherevach I.,
RA Achman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett U.,
RA Felwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
RA Leather S., Moulé S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica."
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640412; CAE34533.1; -.
KW Complete proteome.
SQ SEQUENCE 265 AA; 28715 MW; 4CA85474B566885B CRC64;

Query Match 39.1%; Score 527.5; DB 16; Length 265;
Best Local Similarity 39.6%; Pred. No. 4.6e-26;
Matches 106; Conservative 57; Mismatches 94; Indels 11; Gaps 4;

QY 9 ICALASGIALACGNSQNSNEPAISKTAQTIKVGVMGPEQAVNA-GOVAKKYNLTVE 68
DB 9 IAAFLAGAAVRA-----QPALAD---KPKLISVTGPHAQIFEVKKQEH-AKQGLNTIQ 58
QY 69 LVEFNDYAMPNSAVSKGELDANAMQHKPYLEKDSQEKGLNLTIVGNTFYVPLAGYSTKI 128
DB 59 VIEFSDYQFVVALASGLDLSNSYHQPYLDMAVADRGY-KLVGIAKTAVIFPIGVYSKYV 117
QY 129 KTNLDELKDGATIVPNDPSNLARALITLKEQGLIKDKDNTLSTLDDIVEPKKLYKE 188
DB 118 KTNLDELKDGATIGIPNDPTGGRRALLLQERGLIKAPPEAGLKAFTPIDVVENPKKLFIE 177
QY 189 VDTSAARAIDVDLAVVNNVYAGQVLTASENGVEVEDSPYVNIIVARADKSKAI 248
DB 178 LDAQQLRSUDDTASAVNTNFMLEAGLDPSKDLVRESAESPYANLVVREQDKDADL 237
QY 249 QDFYKAYQTEVEAEAKKQFQDGYIKGM 276
DB 238 RKLVSIVQSAFVKEFILGKYGAIVAAW 265

RESULT 16
Q7W4E3
ID Q7W4E3 PRELIMINARY; PRT; 265 AA.
AC Q7W4E3;
DT 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DE Putative exported protein.
GN BP3723.
OS Bordetella parapertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
NCBI_Taxid=519;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=12822 / ATCC BAA-587;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebailia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Cherevach I.,
RA Achman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett U.,
RA Felwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
RA Leather S., Moulé S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica."
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640412; CAE39006.1; -.

```


KM Complete proteome.
SQ SEQUENCE 265 AA; 28715 MW; 4CA85474B566885B CRC64;

Query Match 38.1%; Score 527.5; DB 16; Length 265;
Best Local Similarity 39.6%; Pred. No. 4,6e-26;

Matches 106; Conservative 57; Mismatches 94; Indels 11; Gaps 4;

QY 9 ICALASGIALAGCSNOSNEPAISKTAAQTIKYGVNAGPEQAVAEVAGVAKKXNLVVE 68
DB 9 LAAPALGAAYFA-----QPALAQD--KPLKIGVAGPHAQIFEVKQEA-AKQGLNIQ 58

QY 69 LVEFNDYAMPNSAVSKGELDANAMQHKPYLEKDSQEKLNNTLVIGNTFVYPLAGYSTKI 128
DB 59 VIEFSDYVQPNVALAGDLDANNSYQHQPYLNNANADRGY-KLVSTAKTVTFPIGVYSKV 117

QY 129 KTLNELKDGATTIAPNDPSNLARALILIEKQGLIKLKDNTNLFSTLLDIVENPKLVIKE 188
DB 118 KTLNELKDGARIGIIPNDPTNGRALLLLEHGLIKLRPAAGKATPIDVENEKRLRFLTE 177

QY 189 VDTSVARAIDVDVLAVNNNNYAGVGLTASENGVEVEDKSPYNNIYARADNKSRAI 248
DB 178 LDAALQPRSLDDTDASAVNTNFMALXGLDPSKDALVRESAESPYANVLVVRBDKDRADL 237

QY 249 QDFVKAQYOTDEVEAEAKKQPKDGVIKGM 276
DB 238 KLVSTIYQSAFVREBFLGKRYKGAIVAAW 265

RESULT 17

QYVVT0 PRELIMINARY; PRT; 265 AA.

ID 07VV70; STRAIN=1021; Pubmed=11481430;

DT 01-OCT-2003 (TREMblrel. 25, Created)

DT 01-OCT-2003 (TREMblrel. 25, Last sequence update)

DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)

DE Putative exported protein.

GN B2818.

OS Bordetella pertussis.

OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;

OC Alcaligenaceae; Bordetella.

OX NCBI_TaxID=520;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;

RX MEDLINE=22827954; Pubmed=12910271;

RA Parkhill J., Sebahthia M., Preston A., Murphy L.D., Thomson N.,

RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,

RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,

RA Achtman M., Atkin R., Baker S., Baaham D., Basson N., Cherevach I.,

RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,

RA Fellwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,

RA Leather S., Mout R., Norbertczak H., O'Neill S., Ormond D., Price C.,

RA Rabinowitsch E., Rutter S., Sanders M., Saunders S., Stevens K.,

RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,

RA Uwin L., Whitehead S., Barrell B.G., Maskell D.J.,

RT "Comparative analysis of the genome sequences of Bordetella pertussis,

RT Bordetella parapertussis and Bordetella bronchiseptica.",

RL Nat. Genet. 35:32-40(2003).

DR EMBL; BX640419; CAE43091.1; -

KM Complete proteome.

SQ SEQUENCE 265 AA; 28743 MW; 6B586474B5668844 CRC64;

Query Match 38.1%; Score 526.5; DB 16; Length 265;

Best Local Similarity 39.6%; Pred. No. 5,3e-26;

Matches 106; Conservative 56; Mismatches 95; Indels 11; Gaps 4;

QY 9 ICALASGIALAGCSNOSNEPAISKTAAQTIKYGVNAGPEQAVAEVAGVAKKXNLVVE 68

DB 9 LAAPALGAAYFA-----QPALAQD--KPLKIGVAGPHAQIFEVKQEA-AKQGLNIQ 58

QY 69 LVEFNDYAMPNSAVSKGELDANAMQHKPYLEKDSQEKLNNTLVIGNTFVYPLAGYSTKI 128

DB 59 VIEFSDYVQPNVALAGDLDANNSYQHQPYLNNANADRGY-KLVSTAKTVTFPIGVYSKV 117

QY 129 KTLNELKDGATTIAPNDPSNLARALILIEKQGLIKLKDNTNLFSTLLDIVENPKLVIKE 188

DB 118 KTLNELKDGARIGIIPNDPTNGRALLLLEHGLIKLRPAAGKATPIDVENEKRLRFLTE 177

QY 189 VDTSVARAIDVDVLAVNNNNYAGVGLTASENGVEVEDKSPYNNIYARADNKSRAI 248

DB 178 LDAALQPRSLDDTDASAVNTNFMALXGLDPSKDALVRESAESPYANVLVVRBDKDRADL 237

QY 249 QDFVKAQYOTDEVEAEAKKQPKDGVIKGM 276

DB 238 KLVSTIYQSAFVREBFLGKRYKGAIVAAW 265

RESULT 18

QYVVT0 PRELIMINARY; PRT; 258 AA.

ID 0921X5; STRAIN=1021; Pubmed=11481430;

DT 01-DEC-2001 (TREMblrel. 19, Created)

DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)

DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)

DE Putative outer membrane lipoprotein transmembrane.

GN R02888 OR SMC03157.

OS Rhizobium meliloti (Sinorhizobium meliloti).

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.

OX NCBI_TaxID=382;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=1021; Pubmed=11481430;

RX MEDLINE=21396507; Pubmed=11481430;

RA Capela D., Barloy-Hubler F., Gouzy J., Boche G., Ampe F., Batut J.,

RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,

RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaire V., Masuy D.,

RA Pohl T., Portetalle D., Puehler A., Purnelle B., Ransperger U.,

RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.,

RT "Analysis of the chromosome sequence of the legume symbiont

RT Sinorhizobium meliloti strain 1021.",

RT Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).

DR EMBL; AL591792; CAC47467.1; -

DR InterPro; IPR004872; Lipoprotein_9.

DR InterPro; IPR004478; YaeC

DR Pfam; PF03180; Lipoprotein_9; 1.

DR TIGRFAWS; TIGR00363; TIGR00363; 1.

KM Complete proteome.

SQ SEQUENCE 258 AA; 27553 MW; 84917D53F26F435A CRC64;

Query Match 37.9%; Score 524.5; DB 16; Length 258;

Best Local Similarity 40.0%; Pred. No. 6,9e-26;

Matches 106; Conservative 57; Mismatches 83; Indels 19; Gaps 3;

QY 12 LASGIALAGCSNOSNEPAISKTAAQTIKYGVNAGPEQAVAEVAGVAKKXNLVVE 71

DB 13 LAAGTAL-----ASTIKGVTFGEHAETMEKKEVAAR-KGLDIEILE 54

QY 72 FNDYAMPNSAVSKGELDANAMQHKPYLEKDSQEKLNNTLVIGNTFVYPLAGYSTKI 131

DB 55 FSDYVVPQALADDLDANNSYQHQPYLNNANADRGY-DIVSGLITITTPGVYSKSKSL 113

QY 132 NELKDGATTIAPNDPSNLARALILIEKQGLIKLKDNTNLFSTLLDIVENPKLVIKEVDT 191

DB 114 DELDGTALIAIPNDPTNGRALLLLEHGLIKLRPAAGKATPIDVENEKRLRFLTE 173

QY 192 SVARAIDVDVLAVNNNNYAGVGLTASENGVEVEDKSPYNNIYARADNKSRAI 251

DB 174 AQLPRSLDDTDASAVNTNFMALXGLDPSKDALVRESAESPYANVLVVRBDKDRADL 233

QY 252 VKAYOTDEVEAEAKKQPKDGVIKGM 276

DB 234 VESTHDKVKAIFYEHRKGLALIPSW 258

RESULT 19

```

08XU08
ID 08XU08 PRELIMINARY; PRT; 529 AA.
AC 08XU08;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Putative OUTERMEMBRANE signal peptide protein.
GN RSC3386 OR R506100.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_Taxid=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM11000.
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoudat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlet M., Billault A., Brothier P., Camus J.C., Catolico L.,
RA Chandler M., Choisme N., Claudel-Renard C., Cumac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Sigrier P., Trebault P., Whalen M., Wincker P., Levy M.,
RA Weisenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL; AL646075; CAD16883.1; -
DR InterPro; IPR004872; Lipoprotein_9.
DR InterPro; IPR004478; YaeC.
DR Pfam; PF03180; Lipoprotein_9; 2.
DR TIGRFAMs; TIGR00363; TIGR00363; 2.
KM Complete proteome.
SQ SEQUENCE 529 AA; 56737 MW; CD6D163194566D17 CRC64;

Query Match 37.4%; Score 517.5; DB 16; Length 529;
Best Local Similarity 39.1%; Pred. No. 56-25;
Matches 108; Conservative 65; Mismatches 84; Indels 19; Gaps 5;

QY 3 FKGINGICALASGIALAGCSNOSNEPAISKTAQTIRKGVNMGPEQAVAEVAGVAKK 62
DB 265 FSRIRARALPALATAIAGAHAEKSP-----IKLGTSGDPAQIWEVYQVAK-K 313
QY 63 YNLTVLVEFNDYAMPNSAVSKGELDANAMQHPYLEKDSQEKGLNNLYIVGNTFVYPLA 122
DB 314 DGLDRIKIFENDYAOQNPALDAGDDPDANGFOHPFDGQVAKGT-KIVNGLTIIVAEWG 372
QY 123 GYSTIKITLNEKDGAITAVPNDPSNLRALILLEKGLIKLK-----DNTNLFSTTLDI 177
DB 373 FYAKKIKSLALKEGKVGIVQNDPSNGRALLLLKAGVITKLKAGAGNGSN--ATPRDV 430
QY 178 VENPKKIVKEVDTSVAARAIDVDLAVVNNNYAGOVGLTASENVFVEDKDSPVNITIV 237
DB 431 VANPKKIKIVELDSQQLPRSLDLDLPAALNTDYAVKNGLTPAKDIALLEDKRGVPANLIA 490
QY 238 AADNKKSKAIDPFYKAYQTDVEAEAKKQFKDGYI 273
DB 491 VAEKDKNQPWKTIVHAYGSEVREKFIETQFKGALL 526

RESULT 20
Q8RFN4 PRELIMINARY; PRT; 261 AA.
AC 08RFN4;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE ABC transporter substrate-binding protein.
GN FN0658.
OS Fuschacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteriia; Fusobacteriales; Fusobacteriaceae;
OC Fusobacterium.
OX NCBI_Taxid=76856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25586;

```

```

RX MEDLINE=21886394; PubMed=11889109;
RA Kapatral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhatnagar A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasileva O., Chu L., Kogan Y., Chaga O., Goldsman E., Bernal A.,
RA Larsen N., D'Souza M., Mainas T., Pusch G., Haselkorn R.,
RA Fongstein M., Kyrides N., Overbeek R.;
RT "Genome sequence and analysis of the oral bacterium Fuschacterium
RT nucleatum strain ATCC 25586.";
RL J. Bacteriol. 184:2005-2018(2002).
DR EMBL; AE010577; AL94854.1; -
DR InterPro; IPR004872; Lipoprotein_9.
DR Pfam; PF03180; Lipoprotein_9; 1.
KM Complete proteome.
SQ SEQUENCE 261 AA; 28768 MW; 60AE194F467F0180 CRC64;

Query Match 37.4%; Score 517, DB 16; Length 261;
Best Local Similarity 41.3%; Pred. No. 21e-25;
Matches 112; Conservative 52; Mismatches 87; Indels 20; Gaps 4;

QY 3 FKGINGICALASGIALAGCSNOSNEPAISKTAQTIRKGVNMGPEQAVAEVAGVAKK 62
DB 7 FGVYGAFLILSAG-ALMG-----TIKQATVPFAHILIKLIPDK-K 47
QY 63 YNLTVLVEFNDYAMPNSAVSKGELDANAMQHPYLEKDSQEKGLNNLYIVGNTFVYPLA 122
DB 48 QGVDLKIVEFDYVTPYALSLDKIEDANFQHKPYLDKFIERRK-LNVSLGNVHVEPLG 106
QY 123 GYSTIKITLNEKDGAITAVPNDPSNLRALILLEKGLIKLKDNNTLFSTTLDIYENPK 182
DB 107 LYSKRIKISINDLKKGDITAIIPDPSNGRALLILHNKGVITLDPKRLFAETPDIVNPK 166
QY 183 KLIVKEVDTSVAARAIDVDLAVVNNNYAGOVGLTASENVFVEDKDSPVNITIVRADN 242
DB 167 KLRKRPVEVQQLRIPLDVTARAIINGVYALQANLSPKDSLIIIEGKSPANILIVRKGD 226
QY 243 KDSKAIDPFYKAYQTDVEAEAKKQFKDGYI 273
DB 227 EKEDIOKTLKALRESEKVKYINEKXSDGSV 257

RESULT 21
Q88DA1 PRELIMINARY; PRT; 284 AA.
AC 088DA1;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Outer membrane lipoprotein.
GN ML4794.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_Taxid=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shlimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AF003005; BAB51370.1; -
DR InterPro; IPR004872; Lipoprotein_9.
DR InterPro; IPR004478; YaeC.
DR Pfam; PF03180; Lipoprotein_9; 1.
DR TIGRFAMs; TIGR00363; TIGR00363; 1.
KM Lipoprotein; Complete proteome.
SQ SEQUENCE 284 AA; 30768 MW; 919CA8873A8C69BF CRC64;

```


Db 180 GGNVLLLDKGVKIKDGTGFKPTVLDITENPKKLEIEVDAAQAAPRALDVAADAAINT 239
 QY 209 NVAGQVGLTASNGVEVEDKDSPPYNIIVARADNKSKAIODFVKAQYQDEVEAEAKQF 268
 Db 240 NVATQGLDPRVQPIILREDEKSPYVNLINRTADKQKPMVKLVDSYHTEPKKEFLVLTGF 299
 QY 269 KQGVVIGW 276
 Db 300 KGAVLPSW 307

RESULT 24

Q9CIN7 PRELIMINARY; PRT; 286 AA.
 AC Q9CIN7
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Outer membrane lipoprotein precursor.
 GN PLPB OR IL0319
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
 OC Bacteria; Firmicutes; Lactobactilales; Streptococcaceae; Lactococcus.
 OX NCBI_TaxID=1360;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=IL1403;
 RX MEDLINE=21335186; PubMed=11337471;
 RA Bolotin A., Winkler P., Manger S., Jallion O., Malarme K.,
 Weissendach J., Ehrlich S.D., Sorokin A.;
 RT "The complete genome sequence of the lactic acid bacterium Lactococcus
 lactis ssp. lactis IL1403."
 RL Genome Res. 11:731-753(2001).
 DR EMBL; AF062659; AKK04417.1; -.
 DR PIR; G86664; G86664.
 DR InterPro; IPR004872; Lipoprotein_9.
 DR Pfam; PF03180; Lipoprotein_9; 1.
 DR Lipoprotein; Complete proteome.
 SQ SEQUENCE 286 AA; 3163 MW; 0C3301DEFA422CDD9 CRC64;

Query Match 36.8%; Score 509.5; DB 16; Length 286;
 Best Local Similarity 40.5%; Pred. No. 7.2e-25;
 Matches 111; Conservative 58; Mismatches 90; Indels 15; Gaps 8;

QY 9 ICALASGIALAGCSNOSNEPAISKTAQTIKYGVNAGPEQ--AAVEAVAGVAKKXVLT 66
 Db 17 IYAIAPLIGF--GQKQAN-----KTVNKTIVKIGIMTGKEDSGIMQVSKTADKRGIT 69
 QY 67 VELVEPNDYAMNSAVSKELDANMOKHPYIEKDSQSEKGLNIVIGNTFVYPLAGYST 126
 Db 70 LKFTHTTYTQNTALKNQDIDLNAFCHYAF--KAMNKAANNGLVAIGDVIYSPISYSK 128
 QY 127 KIKTNELKQATIVAPNDPSVLAARALILLEKQGLIKIKDNTNLFSTLIDIVENPKKIYI 186
 Db 129 QAKNSIDIKEGTIAVPDASVESRALYVKSAGIILKIDVSGQTLATVKDITSNPKNIVT 188
 QY 187 KEVDTSVAAARAIQDVLAVNNNNAQVGLTASNGVEVE--DKDS--PYNIIIVARADNK 243
 Db 189 KELDASQTAARALDSVDAVINNNVAVTAGLKSD--AIFTEPYNKSQSQMINIIVANKDE 247
 QY 244 DSKAIQDFKAYQYQDEVEAEAKQKQF--GVYKGM 276
 Db 248 NNTVYKDVVKAIVEATKTKTIAKAYPDKSTIPAM 281

RESULT 25

Q8FXS4 PRELIMINARY; PRT; 259 AA.
 AC Q8FXS4
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Lipoprotein, YaeC family.
 GN BR2175

OS Brucella suis.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Brucellaceae; Brucella.
 OX NCBI_TaxID=29461;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=1330 / Biovar 1;
 RX MEDLINE=22247741; PubMed=12271122;
 RA Paulsen I.T., Seshadri R., Neilson K.E., Eisen J.A., Heidelberg J.F.,
 Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beanan M.J.,
 Daugherty S.C., Debey R.T., Durkin A.S., Kolonay J.F., Madupu R.,
 Nelson W.C., Ayodeji B., Krahl M., Shetty J., Malek O., Van Aken S.E.,
 Riedmuller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,
 Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;
 RT "The Brucella suis genome reveals fundamental similarities between
 animal and plant pathogens and symbionts."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
 DR EMBL; AE014504; AAN31064.1; -.
 DR TIGR; BR2175; -.
 DR InterPro; IPR004872; Lipoprotein_9.
 DR InterPro; IPR004478; YaeC.
 DR Pfam; PF03180; Lipoprotein_9; 1.
 DR TIGRFAMS; TIGR00363; TIGR00363; 1.
 DR Complete proteome.
 SQ SEQUENCE 259 AA; 27975 MW; A94FA42EB3EBE74B CRC64;

Query Match 36.7%; Score 507; DB 16; Length 259;
 Best Local Similarity 40.0%; Pred. No. 9.1e-25;
 Matches 106; Conservative 51; Mismatches 98; Indels 10; Gaps 3;

QY 12 LASGIALAGCSNOSNEPAISKTAQTIKYGVNAGPEQAAVAVAGVAKKXVLTVEVE 71
 Db 5 LVAAITIALCTAGS-----AFAKTLRVGVTPGPHAQIMEKYEVAARK-GIEIEIOE 55
 QY 72 FNDYAMNSAVSKELDANMOKHPYIEKDSQSEKGLNIVIGNTFVYPLAGYSTIKTL 131
 Db 56 FSDYVNMNALDQALEANCFQKHPYIDNQVADRGF-EIYSVAETVFPMAKTSRKIKTI 114
 QY 132 NELKQATIVAPNDPSVLAARALILLEKQGLIKIKDNTNLFSTLIDIVENPKKIYIEVD 191
 Db 115 AELADGSTIIAIPNDPTGGRALVLADNGILKIDSTGKLKVSADVTEENKIKLFVELDA 174
 QY 192 SVAAARAIQDVLAVNNNNAQVGLTASNGVEVEKDSPPYNIIVARADNKSKAIODF 251
 Db 175 AQLPRESLEDVGAIVNTNVLLEGLDPRKADALVBEQKAPYINILVIRADKADAVKTL 234
 QY 252 VKAYQYQDEVEAEAKQKQGVVIGW 276
 Db 235 IESYHSDEVDQFIATKTFKGSVPAM 259

RESULT 26

Q8YEC4 PRELIMINARY; PRT; 268 AA.
 AC Q8YEC4
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE ABC transporter substrate binding protein.
 GN BME11954.
 OS Brucella melitensis.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Brucellaceae; Brucella.
 OX NCBI_TaxID=29459;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=16W / ATCC 23456 / Biotype 1;
 RX MEDLINE=20020109; PubMed=11756688;
 RA DelVecchio V.G., Kapralov V., Redkar R.J., Patra G., Mujer C., Los T.,
 Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
 Jablonowski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,
 Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,

"The genome sequence of the facultative intracellular pathogen *Brucella melitensis*."; *Proc. Natl. Acad. Sci. U.S.A.* 99:443-448(2002).

DR EMBL; AE009628; AAL5135.1; -.

DR PIR; AD3496; AD3496.

DR InterPro; IPR004872; Lipoprotein_9.

DR Pfam; PF03180; Lipoprotein_9; 1.

DR TrEMBL; TIGR00363; TIGR00363; 1.

KW Complete proteome.

SO SEQUENCE 268 AA; 28948 MW; 1AF605B6B52E1F4B CRC64;

Query Match 36.7%; Score 507; DB 16; Length 268;
Best Local Similarity 40.0%; Pred. No. 9.5e-25;
Matches 106; Conservative 51; Mismatches 96; Indels 10; Gaps 3;

QY 12 LASGALAGCSNOSNEPAISKTAQTIKGVNAGPEQAVAEVAGVAKKNTLTVLVE 71
DB 14 LVAATITLCTAGS-----APAKTLRVGVTGPPHAQIMKVEKVAARK-GIIEIIE 64
QY 72 FNDYAMPNSAVSKGELDANAMQHPYLEKDSOEGLNNLVVGNTPYVPLAGSTKIKTL 131
DB 65 FEDYVMPNVALADGALNANCFQHPYLDNQVADRGF-EIVSAVETVNFPMAMYSRLKTI 123
QY 132 NEIDKQATINVPNDPSNIALRALLLEKQGLIKLKDNTNLFSTLIDIVENPKLVKEVT 191
DB 124 AELADGSTTALPNDPTNGRALLVLADNGLIKLSTKGLVSPADVTENPKIKFEVELDA 183
QY 192 SVAAFAIDVDVLAVANNYAGQVGLTASENGVEFEDKDSRYVNIIVARADNKDSKAIQDF 251
DB 184 AQLPRSLDVGAVINTNTYALAEGLDPRKADALVREGEKAPYNTILVTRADQADNVKTL 243
QY 252 VKAYOTDEVEAEAKKQPFQDGVYKGV 276
DB 244 IESYHSDVKKDPIAKTFKXSVSPAW 268

RESULT 27

088CL5 PRELIMINARY; PRT; 261 AA.

AC 088CL5: MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkman L., Beaman M., Deboy R.T., Daugherty S., Kolonay J.,
RA Madupu R., Nelson W., White O., Peterson J., Knorr H., Hance I.,
RA Chris Lee P., Holtzapple E., Scanlan D., Tran K., Woarnez A.,
RA Utebbeck T., Rizzo M., Lee K., Kosack D., Moestl D., Medler H.,
RA Lander J., Sjepeandic D., Hohnsbeil J., Straetz M., Helm S.,
RA Klewitz C., Eisen J., Timmis K.N., Duesterhoeft A., Tsemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile *Pseudomonas putida* KT2440.";
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AE016793; AAN7030.1; -.
DR TIGR; PP5165; -.
DR InterPro; IPR004872; Lipoprotein_9.
DR Pfam; PF03180; Lipoprotein_9; 1.
KW Complete proteome.
SO SEQUENCE 261 AA; 28159 MW; 9C1B3FE4687FEA13 CRC64;

Query Match 36.6%; Score 506.5; DB 16; Length 261;

Best Local Similarity 41.3%; Pred. No. 9.9e-25;
Matches 109; Conservative 52; Mismatches 88; Indels 15; Gaps 4;

QY 11 ALASGIALAGCSNOSNEPAISKTAQTIKGVNAGPEQAVAEVAGVAKKNTLTVLVE 69
DB 9 ALAAALSFSGIA-----AAEKLVNATPVPHAEIIELEKPLAKSGVLDQIKV 57
QY 70 VERNYAMPNSAVSKGELDANAMQHPYLEKDSOEGLNNLVVGNTPYVPLAGSTKIKTL 129
DB 58 --FTDYVQPNVQDQKLDANVFQTLPYLQNFEGGTHLETWVG-VHPEFGGYSKKVK 114
QY 130 TLNELKQATINVPNDPSNIALRALLLEKQGLIKLKDNTNLFSTLIDIVENPKLVKEV 189
DB 115 ALSLEKGAITVALPNSGNSGRALLLLQKRGVLTLDPRKALATPRDIENPKIKFREL 174
QY 190 DTSVAAIDVDVLAVANNYAGQVGLTASENGVEFEDKDSRYVNIIVARADNKDSKAIQ 249
DB 175 ESANLPLVQLQVDLMDNTNTYALAEGLNPKALVLEGADSPYVNLVAPDNKQSEAIQ 234
QY 250 DFVAYOTDEVEAEAKKQPFQDGVY 273
DB 235 KLAVALSPVKAFAIKKYGAVL 258

RESULT 28

09HT68 PRELIMINARY; PRT; 260 AA.

AC 09HT68: MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Watterer P.,
RA Hickey M.J., Brinkman F.S.L., Huinagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lattig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saiter M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL; AE004963; AAG08890.1; -.
DR PIR; E82957; E82957.
DR CO; GO:0016020; C:membrane; IEA.
DR CO; GO:0004872; F:receptor activity; IEA.
DR CO; GO:0005315; F:transporter activity; IEA.
DR CO; GO:0006810; F:transport; IEA.
DR InterPro; IPR004872; Lipoprotein_9.
DR Pfam; PF03180; Lipoprotein_9; 1.
DR PROSITE; PS00430; TONB DEPENDENT _BEC_1; 1.
KW Receptor; Complete proteome.
SO SEQUENCE 260 AA; 28065 MW; 1E553A0D5C9BEA1A CRC64;

Query Match 36.3%; Score 502.5; DB 16; Length 260;
Best Local Similarity 42.6%; Pred. No. 1.8e-24;
Matches 106; Conservative 52; Mismatches 84; Indels 7; Gaps 3;

QY 29 AAISKTAQTIKGVNAGPEQAVAEVAGV-----AKKNTLTVLVEFNDYAMPNSAVSK 84
DB 12 ALGGLTAAAGSITVAATPVPHAEIIELEKPLAKSGVLDKIK--FTDYVQPNVQVSE 69
QY 85 GELDANAMQHPYLEKDSOEGLNNLVVGNTPYVPLAGSTKIKTLNELKQATINVPN 144

Db 70 KRDLANFQHPYLDDEFNKAG-TDLVAVNGVHIEPLGAVSSKTKLDELPGATVIVN 128

Qy 145 DPNLNARALILEKQGLIKLKNNTLFTTLDIYENPKLVKEVDTSVAARAIIDVDIA 204

Db 129 DATNGGRALILLIDKAGVITKDKNKSITATPKDIVDPKNIKIRELEAATLPVLTQVDVA 188

Qy 205 VVANNAYAGVGLTASENGVEVEDKSPYVNIIVARADNKSADIODPVKAYQIDEVEEA 264

Db 189 LINTNVALEKKNPTDATALIBGSDSPYVNIIVARPDNDSDMQLALALHSAETKQPT 248

Qy 265 KQKQKQGVIT 273

Db 249 QEKYKQAVV 257

RESULT 29

Q8PFO PRELIMINARY; PRT; 269 AA.

AC 08PFO; 01-OCT-2002 (TREMBlrel. 22, Created)

DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE Outer membrane protein.

GN XAC3667.

OS Xanthomonas axonopodis (pv. citri).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;

OC Xanthomonadaceae; Xanthomonas.

OX NCBI_TaxID=92829;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=306 / ATCC 13902 / XV 101;

RX MEDLINE=20222145; PubMed=12024217;

RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,

RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,

RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,

RA Cicarelli R.M.B., Coutinho L.L., Cristina-Santos J.R., El-Dorri H.,

RA Faria J.B., Ferreira A.U.S., Ferreira R.C.C., Ferro M.I.T.,

RA Fournghieri E.F., Franco M.C., Greggio C.C., Gruber A.,

RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,

RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,

RA Martins E.C., Meidanis J., Menck C.F.W., Miyaki C.Y., Moon D.H.,

RA Pereira H.M., Rossi A., Sena U.A.D., Silva C., de Souza R.F.,

RA Spindola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,

RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,

RA Setubal J.C., Kitajima J.P.,

RA "Comparison of the genomes of two Xanthomonas pathogens with differing

RT host specificities."

RI Nature 417:459-463(2002).

RL EMBL: AB012017; AM38510.1; --

DR GO: GO:0016020; C:membrane; IEA.

DR GO: GO:0004872; F:receptor activity; IEA.

DR GO: GO:0005215; F:transporter activity; IEA.

DR GO: GO:0006810; P:transport; IEA.

DR InterPro: IPR004872; Lipoprotein_9.

DR InterPro: IPR000437; Prok_Lipoprote_S.

DR InterPro: IPR000531; TonB_Box.

DR Pfam: PF03180; Lipoprotein_9; 1.

DR PROSITE: PS00013; PROKAR LIPOPROTEIN; 1.

DR PROSITE: PS00430; TONB_DEPENDENT_REC_1; 1.

KW Complete proteome.

SEQUENCE 269 AA; 29097 MW; 75F2ACEB07C2B806A CRC64;

Query Match 36.2%; Score 501; DB 16; Length 269;

Best Local Similarity 41.6%; Pred. No. 2.3e-24;

Matches 109; Conservative 47; Mismatches 98; Indels 8; Gaps 4;

Qy 13 ASGIALAGSNGSNBPAAISKTAQTIKVGWAGPEQAVAEVAG-VAKKYNLTVELVE 71

Db 12 AATLLAAGCGSSG-----SGGGDTLTVAATAVFAHILEVEPLAKQGVKDVAV- 65

Qy 72 FNDVAMPNSAVSKGELDANAMQHKPYLEKDSQEKGLNLTIVVGNTPVYPLAGYSTIKITL 131

Db 66 FNDVQPDQVQKQIDVNYFQTEPPLYLDVAVNRD-KSQLVTVGVHIEPLGAVSSKTKLDELPGATVIVN 124

Qy 132 NELKQATIAVNPDPNSLNARALILEKQGLIKLKNNTLFTTLDIYENPKLVKEVDTSVAARAIIDVDIA 191

Db 125 AELPTGAVAVIPNDPNSNNRALLIDKAGVITKDKNKSITATPKDIVDPKNIKIRELEAATLPVLTQVDVA 184

Qy 192 SVARAIIDVDLAVVNNYAGVGLTASENGVEVEDKSPYVNIIVARADNKSADIODPVKAYQIDEVEEA 251

Db 185 AMLPRIVDQVDLALINTNVYALDAGINPTDATALIBGSDSPYVNIIVARPDNDSDMQLALALHSAETKQPT 244

Qy 252 VKAYQIDEVEEAQKQFQGVIT 273

Db 245 AKALTSPEVKAFIEQKTKGAVV 266

RESULT 30

Q8X85 PRELIMINARY; PRT; 259 AA.

AC 08X85; 01-MAR-2001 (TREMBlrel. 16, Created)

DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)

DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)

DE Hypothetical protein PA3931.

GN PA3931.

OS Pseudomonas aeruginosa.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

OC Pseudomonadaceae; Pseudomonas.

OX NCBI_TaxID=287;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 15692 / PA01;

RX MEDLINE=20437337; PubMed=10984043;

RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,

RA Hickey M.J., Britkhan F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,

RA Garber R.L., Goltzky L., Tolentino E., Westbrook-Wadman S., Yuan Y.,

RA Brody L.L., Coulter S.N., Folger K.R., Kae A., Lardys K., Lim R.M.,

RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,

RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.,

RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an

RT opportunistic pathogen."

RI Nature 406:959-964(2000).

RL EMBL: AB004811; AG07318.1; --

DR PIR: G83153; G83153.

DR InterPro: IPR004872; Lipoprotein_9.

DR Pfam: PF03180; Lipoprotein_9; 1.

KW Hypothetical protein; Complete proteome.

SEQUENCE 259 AA; 28101 MW; C5DF8C298BD233BA CRC64;

Query Match 36.0%; Score 497.5; DB 16; Length 259;

Best Local Similarity 41.4%; Pred. No. 3.7e-24;

Matches 110; Conservative 51; Mismatches 88; Indels 17; Gaps 4;

Qy 9 ICLASIALAGSNGSNBPAAISKTAQTIKVGWAGPEQAVAE-VAGQVAKKYNLTVELVE 67

Db 7 LTLAASVT-----SVSASAALKLVAATPTIPAEILIELIKPTLAKGVLDLQI 53

Qy 68 ELVEFNDVAMPNSAVSKGELDANAMQHKPYLEKDSQEKGLNLTIVVGNTPVYPLAGYSTIK 127

Db 54 KV--FTDYVQPNQVAKKRLDANYPQTLPLETFNFKGKGNLVTVG-VHVEFGGYSKR 110

Qy 128 IKTLLNELKQATIAVNPDPNSLNARALILEKQGLIKLKNNTLFTTLDIYENPKLVK 187

Db 111 YKSLAELPDGATVAIPNEGNSGRALLILQKGLIKLKNNTLFTTLDIYENPKLVK 170

Qy 188 EVLTQVAAAIIDVDLAVVNNYAGVGLTASENGVEVEDKSPYVNIIVARADNKSADIODPVKAYQIDEVEEA 247

Db 171 ELTSALLPRLVDVDDLINTNVYALEKKNPAKDALVLEDRSPYVNIIVARPDNDSDMQLALALHSAETKQPT 230

Qy 248 IODPVKAYQIDEVEEAQKQFQGVIT 273

Db 231 IKTLLAALTSPEVKAFIEQKTKGAVV 256

```

RESULT 31
ID 08P4S9 PRELIMINARY; PRT; 266 AA.
AC 08P4S9;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Outer membrane protein.
GN XCC3628.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Chagallo R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.B.A.,
RA Camarotte G., Camavan F., Cardoso J., Chambergo F., Chapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.U.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi M.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Seta J.A.D., Silva C., de Souza R.F.,
RA Spindola L.A.F., Takita M.A., Tamura R.E., Teixeira B.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:459-463(2002).
DR EMBL; AB012484; AAM42898.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006610; P:transport; IEA.
DR InterPro; IPR004872; Lipoprotein_9.
DR InterPro; IPR000437; Prok_Lipoprot_5.
DR InterPro; IPR000531; TonB_boxC.
DR Pfam; PF03180; Lipoprotein_9; 1.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; 1.
KW Complete proteome.
SQ SEQUENCE 266 AA; 28870 MW; 9023F0E86CDD3E90 CRC64;

Query Match 36.0%; Score 497.5; DB 16; Length 266;
Best Local Similarity 40.1%; Pred. No. 3.8e-24;
Matches 105; Conservative 53; Mismatches 93; Indels 11; Gaps 4;

13 ASGIALAGCNQNSNEPAISKTAQITKVGMAPEQVAEVAQ--VAKEKYLTVELVE 71
12 AATIALALSCGGGGS-----SGGDTITVAATAVPAHEILEVEPLADGVLDRV-- 62
DB
QY FNDYAMNSAVSKGELDANMOKHPYLEKDSQKGLNNTLVIGNTFVYPLAGYSTIKTL 131
72 FNDYAMNSAVSKGELDANMOKHPYLEKDSQKGLNNTLVIGNTFVYPLAGYSTIKTL 131
DB
QY FNDYVQGNDDVQKQIDVNYFQTEPEYIDAVNDR--KSQLVTVGVGHIEPGAVSRFKSL 121
63 FNDYVQGNDDVQKQIDVNYFQTEPEYIDAVNDR--KSQLVTVGVGHIEPGAVSRFKSL 121
DB
QY NELKDQATVAVPNDPSLALILKQGLIKLKDNTNLFSTLIDIVENPKKLVIREVPT 191
132 NELKDQATVAVPNDPSLALILKQGLIKLKDNTNLFSTLIDIVENPKKLVIREVPT 191
DB
QY ADLPFGADVYVLPNDPSNNSRALILDKAGVYIKLADPSNALSQPDIVDPKQIKFELDS 181
122 ADLPFGADVYVLPNDPSNNSRALILDKAGVYIKLADPSNALSQPDIVDPKQIKFELDS 181
DB
QY VAAARAIDVDVLAIVNNVAGVGLTASENGVFVEDKSPYNIIVARADNKSKAIQDF 251
192 VAAARAIDVDVLAIVNNVAGVGLTASENGVFVEDKSPYNIIVARADNKSKAIQDF 251
DB
QY AMLPRVLDQVLDALINTNYALDGLNPTRALAISDSQSYVFLVARADNKDPRVQKL 241
182 AMLPRVLDQVLDALINTNYALDGLNPTRALAISDSQSYVFLVARADNKDPRVQKL 241
DB
QY VKAVQTDVEAEAKKQKQGVYI 273
252 VKAVQTDVEAEAKKQKQGVYI 273
DB
QY AKALTSPVKAIFQKTKGAVL 263
242 AKALTSPVKAIFQKTKGAVL 263

```

```

RESULT 32
ID 088RL7 PRELIMINARY; PRT; 256 AA.
AC 088RL7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ABC transporter, periplasmic binding protein, putative.
GN PPO112.
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Foute D.B., Gill S.R., Pop M., Holmes M.,
RA Brinkac L., Beanan M., Debay R.T., Daugherty S., Kolonay J.,
RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,
RA Chris Lee P., Holtzapple B., Scanlan D., Tran K., Moazzes A.,
RA Uterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,
RA Lauber J., Sjograndic D., Hobeisel J., Straetz M., Helm S.,
RA Kiewitz C., Eisen J., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440.";
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AB016774; AAM65746.1; -
DR TIGR; PPO112; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR004872; Lipoprotein_9.
DR InterPro; IPR000531; TonB_boxC.
DR Pfam; PF03180; Lipoprotein_9; 1.
DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; 1.
KW Complete proteome.
SQ SEQUENCE 256 AA; 27711 MW; 12489C8B6AF86F99 CRC64;

Query Match 35.3%; Score 488; DB 16; Length 256;
Best Local Similarity 42.3%; Pred. No. 1.5e-23;
Matches 105; Conservative 48; Mismatches 89; Indels 6; Gaps 3;

29 AAIKTAQITKVGMAQ--EQVAEVAQVAKKKNLTVELVEFNDYAMPASVSKG 85
9 AAVAFPSKQADTLTVAATPVPAHEILNFKVQLAKEVEELKVK--EFTDIQPNVQVAEK 66
DB
QY ELDNAMOKHPYLEKDSQKGLNNTLVIGNTFVYPLAGYSTIKTLNLDKQATVAVPND 145
86 ELDNAMOKHPYLEKDSQKGLNNTLVIGNTFVYPLAGYSTIKTLNLDKQATVAVPND 145
DB
QY RLDAVFPQHPYLDPEPKAKG--TSLVSVAGVGHIEPLGVSTIKIKKLELSSGATVAVPND 125
67 RLDAVFPQHPYLDPEPKAKG--TSLVSVAGVGHIEPLGVSTIKIKKLELSSGATVAVPND 125
DB
QY PSNLARALILKQGLIKLKDNTNLFSTLIDIVENPKKLVIREVPTVAAARAIDVDVLA 205
146 PSNLARALILKQGLIKLKDNTNLFSTLIDIVENPKKLVIREVPTVAAARAIDVDVLA 205
DB
QY ATNGERALLILDKAGVYIKLKDNTNLFSTLIDIVENPKKLVIREVPTVAAARAIDVDVLA 185
126 ATNGERALLILDKAGVYIKLKDNTNLFSTLIDIVENPKKLVIREVPTVAAARAIDVDVLA 185
DB
QY VNNVYAGVGLTASENGVFVEDKSPYNIIVARADNKSKAIQDFKAVQTDVEAEAK 265
206 VNNVYAGVGLTASENGVFVEDKSPYNIIVARADNKSKAIQDFKAVQTDVEAEAK 265
DB
QY INTVYALEAKINPEKDALATIGSDSPYNIIVARADNKSDMDKGLAALHSPEVKQFI 245
186 INTVYALEAKINPEKDALATIGSDSPYNIIVARADNKSDMDKGLAALHSPEVKQFI 245
DB
QY KQFKDGVYI 273
266 KQFKDGVYI 273
DB
QY EKYKQAVV 253
246 EKYKQAVV 253

RESULT 33
ID 087UN6 PRELIMINARY; PRT; 257 AA.
AC 087UN6;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)

```


DT 01-JUN-2003 (TREMBLrel. 24, last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, last annotation update)
 DE D-methionine-binding lipoprotein MetQ.
 GN METQ-2 OR PSPT05260.
 OS Pseudomonas syringae (pv. tomato).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxId=323;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DC3000;
 RA Buell R., Joardar V., Khouri H., Fedorova N., Tran B., Russell D.,
 RA Berry K., Utterback T., Van Aken S., Feldblum T., Gwin M.,
 RA Dodson R., DeBoy R., Durkin A., Kolonay J., Madupu R., Dougherty S.,
 RA Brinkac L., Beanan M., Haft D., Selengut J., Nelson W., Davidson T.,
 RA White O., Fraser C., Collier A.,
 RA "Complete sequence of Pseudomonas syringae."
 RT Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AB016875; AA058686.1; -
 DR TIGR; PSPT05260; -
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0006101; P:transport; IEA.
 DR InterPro; IPR004872; Lipoprotein_9.
 DR InterPro; IPR000531; TonB boxC.
 DR Pfam; PF03180; Lipoprotein_9; 1.
 DR PROSITE; PS00430; TONB DEPENDENT_REC_1; 1.
 DR Lipoprotein; Complete proteome.
 KM SEQUENCE 257 AA; 27806 MW; 856EE917AB9E10E8 CRC64;

Query Match 35.2%; Score 487; DB 16; Length 257;
 Best Local Similarity 42.9%; Pred. No. 1.7e-23;
 Matches 105; Conservative 45; Mismatches 91; Indels 4; Gaps 3;

QY 30 AISKTAQITIKYVAGPBOQVAE-VAGOVAKEKNLTVELVENDYAMPNSAVKSGELD 88
 DB 13 AISAQNEHTLVAASAVPFAETLEFVKPTLAKGVDLKTV--FNDYIOPNAQVOSQKRLD 70
 QY 89 ANAMQHKPYLEKDSQKGLNIVIGNTFVYPLAGSTIKITKLTNEKOGATAVENDPSN 148
 DB 71 ANFQHPQPLDLENKKG-TDLVAVAKVVEPFGAVSDKFKLDLPGANVALNDATN 129
 QY 149 LRAALILEKQGLIKKONTNLFSTLDIVENPKLVKEVDTSVAARAIDVDLAVNN 208
 DB 130 EGRALLLAKAGILITKDPNSNLSKPSDVNNPKDLKPELEAAATLPVLTQVDLALINT 189
 QY 209 NYAGQVGLTASENGVFEVDKSPYVNIIVARADNDSKAIODFVAYOTDEVEAAKKQF 268
 DB 190 NYALSKLDPTFCALITBGAADSPYANILVTRPDNDSDAIKCLVALDOSPVEKXTLAEKY 249
 QY 269 KDGVI 273
 DB 250 KGAVV 254

RESULT 34
 ID Q7VI95 PRELIMINARY; PRT; 268 AA.
 AC Q7VI95;
 DT 01-OCT-2003 (TREMBLrel. 25, Created)
 DT 01-OCT-2003 (TREMBLrel. 25, last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, last annotation update)
 DE Hypothetical protein yaeC.
 GN YAE C OR H0713.
 OS Helicobacter hepaticus.
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 OC Helicobacteraceae; Helicobacter.
 OX NCBI_TaxId=32025;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 51449 / 3B1;
 MEDLINE=22709201; PubMed=12810954;

RA Suerbaum S., Josephans C., Sterzenbach T., Drescher B., Brandt P.,
 RA Bell W., Droege M., Fattmann B., Fischer H.-P., Ge Z., Hoerster A.,
 RA Holland R., Klein K., Koenig J., Macko L., Mendz G.L., Nyakatura G.,
 RA Schauer D.B., Shen Z., Weber J., Froesch M., Fox J.G.,
 RA "The complete genome sequence of the carcinogenic bacterium
 RT Helicobacter hepaticus."
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7901-7906(2003).
 DR EMBL; AB017146; AAP77310.1; -
 DR Hypothetical protein; Complete proteome.
 KM SEQUENCE 268 AA; 29314 MW; 20948EB2918109F CRC64;

Query Match 35.1%; Score 485; DB 16; Length 268;
 Best Local Similarity 40.0%; Pred. No. 2.4e-23;
 Matches 104; Conservative 56; Mismatches 94; Indels 6; Gaps 5;

QY 15 GIALAGCSNGNEPAISKTAQTI-KYGVVAGPBOQVAEVAQVAKENLTVELVEFN 73
 DB 11 GLVFFGCSDEKN--TSQGTAEKVLKVGATPVFAETLEFI-KPDLKKGIDQIVQFT 67
 QY 74 DYAMPNSAVSKGELDANAMQHKPYLEKDSQKGLNIVIGNTFVYPLAGSTIKITLNE 133
 DB 68 DYTFNVLNDSLDANHOKHPFLDALKDKQGL-NLESIASIHIEPLGFYSHKFKIDS 126
 QY 134 LKDGATIAVPNDPSNLARALILEKQGLIKKONTNLFSTLDIVENPKLVKEVDTSV 193
 DB 127 LPQGTIAIPDPSNGRALLLDSKGIITLADTNINLNEIDIVENPKIKIPVEAAL 186
 QY 194 AARAIDVDLAVANNYAGVGLTASENGVFEVDKSPYVNIIVARADNDSKAIODFVK 253
 DB 187 LPRITDVGAVINGNALQGLKSD-ALFLESGQSPYANILVOSTRVADENILQTLKN 245
 QY 254 AYOTDEVEAAKKQFQKQVI 273
 DB 246 ALQSQKXKDFIQHYQGEIV 265

RESULT 35
 ID Q832Y8 PRELIMINARY; PRT; 277 AA.
 AC Q832Y8;
 DT 01-JUN-2003 (TREMBLrel. 24, Created)
 DT 01-OCT-2003 (TREMBLrel. 24, last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, last annotation update)
 DE Lipoprotein, yaeC family.
 GN EF2080.
 OS Enterococcus faecalis (Streptococcus faecalis).
 OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
 OX NCBI_TaxId=151;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=V583 / ATCC 700802;
 RC MEDLINE=22550857; PubMed=12663927;
 RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,
 RA Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
 RA Tettelin H., Dodson R.J., Umayam L., Brinkac L., Beanan M.,
 RA Dougherty S., DeBoy R.T., Durkin S., Kolonay J., Madupu R., Nelson W.,
 RA Vanathaveyan J., Tran B., Upton J., Hansen T., Shetty J., Khouri H.,
 RA Utterback T., Radune D., Ketchum K.A., Dougherty B.A., Fraser C.M.,
 RA "Role of mobile DNA in the evolution of vancomycin-resistant
 RT Enterococcus faecalis."
 RL Science 299:2071-2074(2003).
 DR EMBL; AB016953; AA081813.1; -
 DR TIGR; EF2080; -
 DR InterPro; IPR004872; Lipoprotein_9.
 DR InterPro; IPR000437; Prok_lipoprot_5.
 DR Pfam; PF03180; Lipoprotein_9; 1.
 DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.
 DR Lipoprotein; Complete proteome.
 KM SEQUENCE 277 AA; 30445 MW; 0A64B6315479551 CRC64;

Query Match 34.8%; Score 481; DB 16; Length 277;
 Best Local Similarity 36.8%; Pred. No. 4.5e-23;
 Matches 102; Conservative 62; Mismatches 101; Indels 12; Gaps 5;

```

QY 3 FKGINGICLALAGCNSGNSNEPAISKTAAQITKVGWAGPEQAVAEVAGVAK 62
DB 4 FSKTILILVLA-FTIAGCAGS-----VKDTKETVTKGVGTGRNDEWESVKRLKK- 56
QY 63 YNLTVELVEFNDYAMPNSAVSKGELDANAMOKHPYLEKDSQEKGINLTVIGNTFVYPLA 122
DB 57 -NIDLOLVEFTDYTGNNALAEKEITLNAFQHOIFLDNNKHHG--TKVVISGNTVNAFLG 114
QY 123 GYSTKIKTINELKDGATIAVPNDPSNLARALILLEKQGLIKLKDNTNLFSTLLD-VEHPK 182
DB 115 IYANKLKIDITIKIDGGEIAPNDPTNGRALILQLTQRAIKVDPAKQOLPTVSDITENKR 174
QY 183 KVIKKEVDTSVAAARIDVDLAVNNNNYAGQVGLTASENGVEVE--DKDSPYNIIVAR 239
DB 175 QLKITLIDATQTKRALQDVPDASVINGMAYDAGYTPDKRAIFLEPNEKAKRYNIIVAR 234
QY 240 ADNKDSKALQDFVKAQYQDVEAEAKKQFKQGVIKGW 276
DB 235 EEDQENKLVQKVVEEYQOEETKKVIAETSKGANVPAM 271

RESULT 36
Q9PPE7 PRELIMINARY; PRT; 257 AA.
ID Q9PPE7
AC Q9PPE7
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Putative periplasmic protein.
GN C00772C.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Campylobacteraceae; Campylobacter.
CX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 11168;
RA MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Kelsey J.M., Churcher C.,
  Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtrop S.,
  Jorgensen K., Karlyshev A.V., Moule S., Pallen C.W., Penn C.W.,
  Rasmussen M.A., Rutherford K.M., van Vliet A.H.M.,
  Whitehead S., Barrall B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
  reveals hypervariable sequences."
RT Nature 403:665-668(2000).
RL EMBL: AL139076; CAB73037.1; -.
DR EMBL; D81348; D81348.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0005215; F:transporter activity; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR004872; Lipoprotein_9.
DR InterPro: IPR000531; TonB box.
DR Pfam: PF03180; Lipoprotein_9; 1.
DR PROSITE: PS00430; TONB_DEPENDENT_REC_1; 1.
DR Complete proteome.
SQ SEQUENCE 257 AA; 28638 MW; D03782C849297381 CRC64;

Query Match 34.5%; Score 477.5; DB 16; Length 257;
Best Local Similarity 41.5%; Pred. No. 6.9e-23;
Matches 100; Conservative 50; Mismatches 78; Indels 13; Gaps 5;

QY 30 AISKTAQITKVGWAGPEQAVAEVAGVAK--EKNTLVELVEFNDYAMPNSAVSKGEL 87
DB 15 SLNNALETTIVATPVPH--ABILEQVPRDLKQKQKLEIKETFDYVLFNLAVNGSA 71
QY 88 DANMOKHPYLEKDSQEKGINLTVIGNTFVYPLAGSTKIKTINELKDGATIAVNDPS 147
DB 72 DANFQITVLEENKKG--TKLIKVAIIEPMAYSKKYSLLDIDKEGVKAIANDPT 130
QY 148 NLARALILEKQGLIKLKDNTNLFSTLLD-VEHPKLVKEVDTSVAAARIDVDLAVN 207

```

```

DB 131 NESRPLDIIAKKGLVFKDKA--LKTPLDIDPKKIKFEVLEKPAQLPRALNDVDFAVIN 188
QY 208 NNTVAGVGLTASENGVEVEKDSPPYNIIVARNDNSKALQDFVKAQYQDVEAEAKKQ 267
DB 189 SNVALSNLNPADDSVFTEDEKSPYANILVVRGHNDEPKIKALIQLOSDKI-----KQ 243
QY 268 F 268
DB 244 F 244

RESULT 37
Q9BEL4 PRELIMINARY; PRT; 273 AA.
ID Q9BEL4
AC Q9BEL4
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical conserved protein.
GN O83327.
OS Oceanobacillus theysensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.
CX NCBI_TaxID=182710;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HTE831 / DSM 14371 / JCM 11309;
RA MEDLINE=2220767; PubMed=12235376;
RA Takami H., Takaki Y., Uchiyama I.;
RT "Genome sequence of Oceanobacillus theysensis isolated from the Iheya
  Ridge and its unexpected adaptive capabilities to extreme
  environments."
RT Nucleic Acids Res. 30:3927-3935(2002).
RL EMBL; AP004604; BAC15283.1; -.
DR InterPro: IPR004872; Lipoprotein_9.
DR InterPro: IPR000437; Prok Lipoprot_5.
DR Pfam: PF03180; Lipoprotein_9; 1.
DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 273 AA; 29904 MW; B05CB0F7A78B211F CRC64;

Query Match 33.9%; Score 469.5; DB 16; Length 273;
Best Local Similarity 34.8%; Pred. No. 2.4e-22;
Matches 96; Conservative 66; Mismatches 111; Indels 3; Gaps 3;

QY 1 MFKINGICLALAGCNSGNSNEPAISKTAAQITKVGWAGPEQAVAEVAGVAK 60
DB 1 MKKGLTAIVFLVAL-LAAGSSESSGASEGLYDEKLVIGTAGHHEILEKVTLEAE 59
QY 61 EKNLTVELVEFNDYAMPNSAVSKGELDANAMOKHPYLEKDSQEKGINLTVIGNTFVY 120
DB 60 EE-GINIEETVFTDYVMPNVALDGEIDLNI FQTEPYFNAIREDRNL-DIVKSFDTVP 117
QY 121 LAGSTKIKTINELKDGATIAVPNDPSNLARALILLEKQGLIKLKDNTNLFSTLLDVEN 180
DB 118 MGIYSLDVGVSELPREGATIGLPDPTNSGRALLLFEQAGLIEIDPFTGINSIVRDIEN 177
QY 181 PKLVKKEVDTSVAAARIDVDLAVNNNNYAGQVGLTASENGVEVEKDSPPYNIIVARA 240
DB 178 NGNIEFVLEDSAQARQOLEIIDAALNTNPAILEGFTPTDAIFIEKDSPPYINHAARA 237
QY 241 DNKDSKALQDFVKAQYQDVEAEAKKQFKQGVIKGW 276
DB 238 ENEDDEITQKLADIYRSEEVTFIEEBFGSLVPSW 273

RESULT 38
Q9PPE8 PRELIMINARY; PRT; 256 AA.
ID Q9PPE8
AC Q9PPE8
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)

```

DE Putative periplasmic protein.
GN C00771C.
RA Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCCTC 11168;
RX MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Feltham T., Holtroyd S.,
RA Jagers K., Karlyshev A.V., Moule S., Fallen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrett B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences."
RL Nature 403:665-668(2000)
RU EMBL: AL139076; CAB73036.1; -
DR PIR: C81348; C81348.
DR InterPro: IPR004872; Lipoprotein_9.
DR Pfam: PF03180; Lipoprotein_9; 1.
KW Complete proteome.
SQ SEQUENCE 256 AA; 28843 MW; B64BA955ED452470 CRC64;
Query Match 33.8%; Score 468; DB 16; Length 256;
Best Local Similarity 42.2%; Pred. No. 2.8e-22;
Matches 103; Conservative 52; Mismatches 83; Indels 6; Gaps 4;
QY 31 ISKTAQTIKVGMAPEQVAEAGVAKEX-YNLTVLVEFNDYAMPNSAVSKGLDANMCKPYL 89
DB 15 LSLFQNTITGATPFGSLHMDPRKNGYEL-KIVERSDYILNRLBEKELDA 72
QY 90 NAMQHPKYLEKDSQSEKGLNNLVIVGNTFVYPLAGYSTKIKTLNELKDGATIAVPNDPSNL 149
DB 73 NLYQHPKPLEEYVNLKKG-SNLVATTPVLLAPGVYSKIKLNLNLEKAGRAVAPNDATNE 131
QY 150 ARALLILEQGLIKKDNNTLFTTLDIYENPKKVIKEVDNSVAAARADVDVLAIVNNN 209
DB 132 SRLLELEKAKLEL--NKNTLKTPIIDINKPKKLEFILLKAAQLFRALDDVDLAINSN 189
QY 210 YAQVGLTASENGFEVDKDSPPYNIIVARADNKSKAIDPFYKAYQTDVEVEAKKQPK 269
DB 190 FALGAGLNSKDTIFEDKNSPVYVYVVRSEKSKSEKTKVIDELIRSKFAIINENHX 249
QY 270 DGYI 273
DB 250 DILI 253
RESULT 39
O26084 PRELIMINARY; PRT; 271 AA.
AC O26084;
DT 01-JUN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Outer membrane protein.
GN HP1564.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klein H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,

RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Matthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori."
RL Nature 388:539-547(1997).
DR EMBL: AE000654; AAD08604.1; -
DR PIR: D64715; D64715.
DR TIGR: HP1564; -
DR InterPro: IPR004872; Lipoprotein_9.
DR Pfam: PF03180; Lipoprotein_9; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 271 AA; 30151 MW; 23786AEFB5C6CB5B CRC64;
Query Match 33.8%; Score 467.5; DB 16; Length 271;
Best Local Similarity 42.1%; Pred. No. 3.2e-22;
Matches 99; Conservative 46; Mismatches 85; Indels 5; Gaps 3;
QY 39 IKGVMAPEQVAEAGVAKEXYNTLTVLVEFNDYAMPNSAVSKGLDANMCKPYL 98
DB 39 LKVGANPVFPAQIIQSVDDLKEX-GIKVTVSFTDYLPLALINDSLDANVFQHRPYL 97
QY 99 EKDSQEKGLNNLVIVGNTFVYPLAGYSTKIKTLNELKDGATIAVPNDPSNLARALLLEK 158
DB 98 DRFNDKRM-HVGLANIHVEPLRFYSOKITDINKLKKGSYIAVPNDPANOGRLILLHK 156
QY 159 QGLIKLKDNTLFTTLDIYENPKKVIKEVDNSVAAARADVDVLAIVNNNVAQVGLTA 218
DB 157 QGLIALKQPSNLVTEFPIVGNPNYIKIKPLEALLPKVLGVDGAILITGVYLAQKLTG 216
QY 219 SENGFEVDKDSPPYNIIVARADNKSKAIDPFYKAYQTDVEVEAKKQPKDGYI 273
DB 217 A--LFESDKDSFYANLIAAREDNAQDEALKTLEALQSEKTRKFLIDTYGAIL 268
RESULT 40
O26084 PRELIMINARY; PRT; 266 AA.
AC O26084;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Outer membrane lipoprotein precursor.
GN PLPC OR IL0320.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacilliales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Wincker P., Manger S., Jallion O., Malarne K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403."
RL Genome Res. 11:731-753(2001).
DR EMBL: AE006269; AAK04418.1; -
DR PIR: H86664; H86664.
DR InterPro: IPR004872; Lipoprotein_9.
DR Pfam: PF03180; Lipoprotein_9; 1.
KW Lipoprotein; Complete proteome.
SQ SEQUENCE 266 AA; 31554 MW; 398725DC477A4D65 CRC64;
Query Match 33.6%; Score 464.5; DB 16; Length 266;
Best Local Similarity 40.5%; Pred. No. 5.3e-22;
Matches 113; Conservative 54; Mismatches 101; Indels 11; Gaps 7;
QY 1 MNFKINGICALAGIAGCSNOS-NEPAISKTAQTIKVGMAE--EQAVAEVAGQ 57
DB 1 MNPKNRNIYALAVIYALVAFSLNHQASKSTAGKTAVGVGIMSGDKODQEWKSVAK 60
QY 58 VAKKYNLTVLVEFNDYAMPNSAVSKGLDANMCKPYLEKDSQSEKGLNNLVIVGNTF 117

Db 61 TAKEKYDLKLFYFSPDYNQNEALLSGDIDVNAFOSYNYV-KTMNKAHKSIDIYAVGNTY 119
QY 118 VYPLAGYSTKIKTNLNEKOGATIAVPNDPSNLARALILEKQGLIKL--KDNNTLFE STL 175
Db 120 ITPWHIYSKEVSLSLSDKEGSTVAIPNDASNEBRALFVLOSAGLIKLTTSDSKLVGLP- 178
QY 176 DIVENPKCLVIVKEVDTSVAAARAIIDVDLAVVNNNVAQVGLTASENGVE--DKDSPY 232
Db 179 DITENPHQLKFKEXVDASQTPRALDSVALSVNNTNYATAASLPNSES-VYMEPLNKTSAGY 237
QY 233 VNIIVAAADNKDSKAIODFYKAYOTDEVEAEAKQFKDG 271
Db 238 INFIAATSKKEKNKVVYKEVAKAYASKATEKAIKEQIFPDG 276

Search completed: June 16, 2004, 11:08:08
job time : 49 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 16, 2004, 11:05:36 ; Search time 22 Seconds
(without alignments)
647,671 Million cell updates/sec

Title: US-10-018-672-2
Perfect score: 1383
Sequence: 1 MNFKINGICALASGIALAG.....TDEVEAKKQPKDGVIXGW 276

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/1aa/5A COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/beck11les1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	775.5	56.1	289	US-09-489-039A-12337	Sequence 12337, A
2	757.5	54.8	279	US-09-543-681A-6534	Sequence 6534, Ap
3	685.5	49.6	285	US-09-489-039A-10641	Sequence 10641, A
4	671.5	48.6	279	US-09-543-681A-4548	Sequence 4548, Ap
5	651	47.1	272	US-09-489-039A-10410	Sequence 10410, A
6	596	43.1	279	US-09-543-681A-6914	Sequence 6914, Ap
7	520	37.6	118	US-09-540-236-2991	Sequence 2991, Ap
8	503.5	36.4	305	US-09-252-991A-18873	Sequence 18873, A
9	494.5	35.8	271	US-09-252-991A-20200	Sequence 20200, A
10	481	34.8	277	US-09-071-035-30	Sequence 30, App1
11	481	34.8	278	US-09-134-000C-4965	Sequence 4965, Ap
12	475	34.3	269	US-09-543-681A-4774	Sequence 4774, Ap
13	462	33.4	257	US-09-071-035-32	Sequence 32, App1
14	449.5	32.5	288	US-09-107-532A-7234	Sequence 7234, Ap
15	441	31.9	291	US-09-134-001C-4005	Sequence 4005, Ap
16	416	30.1	280	US-08-956-171E-5195	Sequence 5195, Ap
17	373.5	27.0	279	US-09-134-001C-4878	Sequence 4878, Ap
18	366	26.5	269	US-09-107-532A-6443	Sequence 6443, Ap
19	364	26.3	296	US-09-328-352-6924	Sequence 6924, Ap
20	364	26.3	326	US-09-489-039A-9512	Sequence 9512, Ap
21	358.5	25.9	261	US-09-634-238-420	Sequence 420, App
22	356	25.7	260	US-08-961-083-32	Sequence 32, App1
23	356	25.7	260	US-09-536-784-32	Sequence 32, App1
24	346	25.0	273	US-08-956-171E-5196	Sequence 5196, Ap
25	337.5	24.4	286	US-09-328-352-5702	Sequence 5702, Ap
26	335	24.2	304	US-09-328-352-7028	Sequence 7028, Ap
27	325.5	23.5	294	US-09-328-352-5549	Sequence 5549, Ap

28	324	23.4	272	US-09-071-035-22	Sequence 22, App1
29	322	22.3	342	US-09-252-991A-25420	Sequence 25420, A
30	321.5	22.2	280	US-09-489-039A-13968	Sequence 13968, A
31	319.5	23.1	220	US-09-634-238-283	Sequence 283, App
32	318	23.0	251	US-09-071-035-24	Sequence 24, App1
33	315.5	22.8	272	US-09-071-035-10	Sequence 10, App1
34	308.5	22.3	253	US-09-071-035-12	Sequence 12, App1
35	301.5	21.8	278	US-09-134-000C-5006	Sequence 5006, Ap
36	257.5	18.6	275	US-09-198-452A-291	Sequence 291, App
37	209	15.1	306	US-09-489-039A-10023	Sequence 10023, A
38	108	7.8	91	US-09-134-000C-3687	Sequence 3687, Ap
39	106	7.7	347	US-09-328-352-6764	Sequence 6764, Ap
40	105	7.6	794	US-09-134-000C-5518	Sequence 5518, Ap
41	103	7.4	503	US-09-071-035-360	Sequence 360, App
42	103	7.4	1074	US-09-071-035-358	Sequence 358, App
43	103	7.4	1074	US-09-071-035-394	Sequence 394, App
44	103	7.4	1096	US-09-134-000C-5764	Sequence 5764, Ap
45	102.5	7.4	308	US-09-134-001C-4400	Sequence 4400, Ap

ALIGNMENTS

RESULT 1
US-09-489-039A-12337
; Sequence 12337, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709, 2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIORITY FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12337
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12337

Query Match 56.1%; Score 775.5; DB 4; Length 289;
Best Local Similarity 55.2%; Pred. No. 3.3e-73;
Matches 153; Conservative 45; Mismatches 68; Indels 11; Gaps 3;

QY 2 MNFKINGICALASGIALAGCSNQSNEPAAISTTAQTIKGVWAGPEQAVAVAGQVAK 61
DB 22 NFKTPAVAGALIGSLALVGCQDEKDP-----NHIKGVYVAGFQVAVAGQVAK 73
QY 62 KYNLIVELVEFNDYAMPNSAVSKGELDANAWQHRYLEKDSQEKLANLIVYGNTPVPL 121
DB 74 KIGLDVEIVTDFNDYLPENALSKGIDIVNAFOHKKPYLLDOQIKDRGY-KLIVAGNTFVYPI 132
QY 122 AGSTTKTLMNLEKOGATFVNDPSNLARAILLEKQGLIKLKNNTNLFSTLIVENP 181
DB 133 AGSKTKISLDELQPSQIAVENDPTNCRSLLLAQGVGLKKGVGILLPTSLDIVENP 192
QY 182 KKLIVKEVTSVAARAIDP--VDLAVVNNNVAGQVGLTASENGVEVEDKDSFYNIIVAR 239
DB 193 KKLIVKEVTSVAARAIDP--VDLAVVNNNVAGQVGLTASENGVEVEDKDSFYNIIVAR 252
QY 240 ADNKSALQDPYKAYQCTDEVEAEAKKQPKDGVIXGW 276
DB 253 EDNKAENYKRFVQAVQSDVEVEAEANKIFNGAVXGW 289

RESULT 2
US-09-543-681A-6534
; Sequence 6534, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:

APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR APPLICATION NUMBER: 2000-04-05
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO: 6534
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-6534

Query Match 54.8%; Score 757.5; DB 4; Length 279;
Best Local Similarity 52.7%; Pred. No. 2.4e-71;
Matches 147; Conservative 49; Mismatches 70; Indels 13; Gaps 4;

QY 1 MFPERKINGICALAGSIALAGCSNOSNEPAISKTAQTIKVGWAGGEOVAEYAGGVAK 60
DB 11 LKPKSLAVSALVGLALAGGCEKEKDP-----NHIRGVISGSEQOVAEYAKGVAK 62
QY 61 EKVNTVLEFNDYAMPNSAVSKGELDANAMQHKPYLEKDSQKGLN-ILVIVGNTFVY 119
DB 63 DRYGLNVLVTFNDFVMPNESLSRGDIDINAFQHKPYL--DQQTIDNNYKITAAGNTFY 120
QY 120 PLAGYSTIKITLNEIKDGAITAVPNDPSNLARALLLEKQGLIKLKDNTNLFSTLTDIVE 179
DB 121 PLAGYSKKITPDLADLPDGAQVAILPNDPTNIGRSILILEKYGVLVLEKGVGLPTKDIIE 180
QY 180 NPKVLVIEVDTSVAARAID--VDLAVNNNNYAGVGLTASENGVEVEKDSFYNNITV 237
DB 181 NPKVQLVELEAPQLPFSLDQKLYLAVINTTASQVNLTPAKGKGFVEDKDSFYNNITV 240
QY 238 ARADNKSKAIDPFVKAQVQTEVEAEAKKQFKDGVIKGW 276
DB 241 AREDNKODENKKEFTIQSYQIDEVDSANKTFNGGAVVGMW 279

RESULT 3
US-09-489-039A-10641
; Sequence 10641, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10641
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10641

Query Match 49.6%; Score 685.5; DB 4; Length 285;
Best Local Similarity 50.6%; Pred. No. 9.4e-64;
Matches 135; Conservative 53; Mismatches 68; Indels 11; Gaps 3;

QY 12 LASGIALAGCSNOSNEPAISKTAQTIKVGWAGGEOVAEYAGGVAKKXNLTVEIVE 71
DB 28 MLALGALVGDQKGN-----AKHIVGVINGAEQDAEYAKKVAEKIGLEVLVG 79
QY 72 FNDYAMPNSAVSKGELDANAMQHKPYLEKDSQKGLNINIVGNTFVYPLAGYSTIKITL 131
DB 80 FSGSLPNESTNAGDDIDANVQHRPFLQDNKAHNT-HLVAAGNTFVFPAGYSRKITSV 138
QY 132 NELKDGATIAVPNDPSNLARALLLEKQGLIKLKDNTNLFSTLTDIVEPKLVIEVD 191

DB 139 AELKDGATIAVPNDPSNLARALLLEKQGLIKLKDNTNLFSTLTDIVEPKLVIEVD 198
QY 192 SVARAID--VDLAVNNNNYAGVGLTASENGVEVEKDSFYNNITVARADNKSQAIQ 249
DB 199 AQLPFLVLDPKVDVAIIISTTYIQGTGLSPVRDGFIEDKNSPYNNITVREDNKAENVK 258
QY 250 DFKAYQTEVEAEAKKQFKDGVIKGW 276
DB 259 EFWQSTQSPPEVAKAETTFNGGAVVGMW 285

RESULT 4
US-09-543-681A-4548
; Sequence 4548, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4548
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-4548

Query Match 48.6%; Score 671.5; DB 4; Length 279;
Best Local Similarity 50.6%; Pred. No. 2.7e-62;
Matches 135; Conservative 49; Mismatches 72; Indels 11; Gaps 3;

QY 12 LASGIALAGCSNOSNEPAISKTAQTIKVGWAGGEOVAEYAGGVAKKXNLTVEIVE 71
DB 22 LLAGTLLTGCDNNNNRYA-----IKVGVINGAEQDAEYAKKVAEKIGLEVLVG 73
QY 72 FNDYAMPNSAVSKGELDANAMQHKPYLEKDSQKGLNINIVGNTFVYPLAGYSTIKITL 131
DB 74 FSGSLPNDPPLANKELDANVQHRPFLQDNRSRGY-ILVAVGNTFVFPAGYSTIKIKHP 132
QY 132 NELKDGATIAVPNDPSNLARALLLEKQGLIKLKDNTNLFSTLTDIVEPKLVIEVD 191
DB 133 QELKGDITIALPNDPTNIGRALLLLEKQGLITLNPDSGLPTLTDIDNPDKIMOLEG 192
QY 192 SVARAID--VDLAVNNNNYAGVGLTASENGVEVEKDSFYNNITVARADNKSQAIQ 249
DB 193 AQLPFLVLDPKVDVAIIISTTYIQGTGLSPVRDGFIEDKNSPYNNITVREDNKAENVK 252
QY 250 DFKAYQTEVEAEAKKQFKDGVIKGW 276
DB 253 DFKAYQSPPEVATAETTFKGAIQGW 279

RESULT 5
US-09-489-039A-10410
; Sequence 10410, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10410
; LENGTH: 272

Tue Jun 22 11:04:04 2004

us-10-018-672-2.ra1

Page 3

TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10410

Query Match 47.1%; Score 651; DB 4; Length 272;
Best Local Similarity 46.4%; Pred. No. 3.7e-60;
Matches 128; Conservative 59; Mismatches 73; Indels 16; Gaps 5;

5 KINGICALASGIA-LAGGS-NOSNEPAISKTAAQTIKVGWAGPEQAVAEVAGQVAKK 62
9 KLFSLALATSVALLSACSPEBDNK-----VVAINTGPDEAIKVAQVAKDK 57
63 YNLVLEVEFNDYAMPNSAVSKGELDANAMQHRYLEKDSQEKGLNVLIVGNTFYVPLA 122
58 YHLDEEVVSFNDYVLPNEALNNKQVDANAFQTLFYLEAQSKERGY-KFAVVGKTFVPEIA 116
123 GYSTKIKTLNELKOGATIAVNDPSNTARALILEKOGILKDKNTNLFSTLIDIVENPK 182
117 AYSHRIKISELPBGATVITISNETTTLSRLSLLLQAGGLKLRKGVGYLPTLIDIDNPK 176
183 KLVIKEVDTSVAAARAIID--VDLAVNNNNVAGVGLTASNGVFEVDKDSFYNNIIVARA 240
177 QKIYEVDTLPOLTRLDLPNVLSTIINNPFSAQGLSARQGLFMEGDSFYVAAIYARE 236
241 DNKSKAIQDPEVKAQVOTDEVEAEAKKQKFDGVIKGM 276
237 DNKSKRIQELKAAFTQSEVAEKAKEYKGDIAIKGM 272

RESULT 6

US-09-543-681A-6914
Sequence 6914, Application US/09543681A
Patent No. 6605709

GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
FILE REFERENCE: 2709 1002-001
CURRENT APPLICATION NUMBER: US/09/543, 681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 6914
LENGTH: 279
TYPE: PRT
ORGANISM: Proteus mirabilis
US-09-543-681A-6914

Query Match 43.1%; Score 596; DB 4; Length 279;
Best Local Similarity 42.1%; Pred. No. 2.4e-54;
Matches 112; Conservative 62; Mismatches 78; Indels 14; Gaps 4;

14 SGIALAGGS-NOSNEPAISKTAAQTIKVGWAGPEQAVAEVAGQVAKKYNLVELVER 72
25 TSLVLTACSPDEGKRP-----LKVAINTPDQIMDEVVKKLKEKGLDIDKVTIF 74
73 NDYAMPNSAVSKGELDANAMQHRYLEKDSQEKGLNVLIVGNTFYVPLAGYSTRKIKTLN 132
75 NDYVLPNEALNNKQVDANAFQTLFYLEAQSKERGY-QFEIIGKTFIPPIAAVSKRIKTIIE 133
133 ELKOGATIAVNDPSNTARALILEKOGILKDKNTNLFSTLIDIVENPKKLVKEVDT 182
134 ALPQATVAAISNEATTLGRSLILLQAGGLIKLRKGVGYLPTLIDIDNPKKRAEIDTP 193
193 VAAARAIID--VDLAVNNNNVAGVGLTASNGVFEVDKDSFYNNIIVARA DNKSKAIQD 250
194 QLTSLSPNTYLSIINNPFSSQGLLAKRGDLFMENTDSFYNNIIVARA IDKDNERLKK 253
251 FVKAQVOTDEVEAEAKKQKFDGVIKGM 276
254 LVAVFQSDIELQKAGVYKGDIAIKGM 279

RESULT 7

US-09-540-236-2991
Sequence 2991, Application US/09540236
Patent No. 6673910

GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRI
FILE REFERENCE: 2205 2005-001
CURRENT APPLICATION NUMBER: US/09/540, 236
CURRENT FILING DATE: 2000-04-04
NUMBER OF SEQ ID NOS: 3840
SEQ ID NO 2991
LENGTH: 118
TYPE: PRT
ORGANISM: M. catarrhalis
US-09-540-236-2991

Query Match 37.6%; Score 520; DB 4; Length 118;
Best Local Similarity 36.3%; Pred. No. 5.8e-47;
Matches 104; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

1 MNGKINGICALASGIALAGGSNOSNEPAISKTAAQTIKVGWAGPEQAVAEVAGQVAK 60
11 MNGKIFGICALASGIALAGGSNOSNEPAISDTAAQTIKVGWAGPEQAVAEVAGQVAK 70
71 EKNLTVLEVEFNDYAMPNSAVSKGELDANAMQHRYLEKDSQEKGLN 118

RESULT 8

US-09-252-991A-18873
Sequence 18873, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252, 991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 18873
LENGTH: 306
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18873

Query Match 36.4%; Score 503.5; DB 4; Length 306;
Best Local Similarity 43.0%; Pred. No. 1.5e-44;
Matches 107; Conservative 51; Mismatches 84; Indels 7; Gaps 3;

29 AAIKTAQTIKVGWAGPEQAVAEVAGV---AKKYNLTVLEVEFNDYAMPNSAVSK 84
58 AALGLTAAQASLTVAATPVPFAHILNVKKLKEGVDLTK--EFTDYVQPNVQVSE 115
85 GEDANAMQHRYLEKDSQEKGLNVLIVGNTFYVPLAGYSTRKIKTLNELKOGATIAVNP 144
116 KRLDANFPQHPRYLDFBNKAKG-TDLVAVTGVHIEPLAGYSSKTKLDELBSGATIVIN 174
145 DPSNTARALILEKOGILKDKNTNLFSTLIDIVENPKKLVKEVDTSVAAARAIIDVDVIA 204
175 DATNGRALLILLDKAGVILKDKNKSITATPKDIDVNPKNIKIRLEDAATLPPVLTVQVMA 234
205 VVNNNNVAGVGLTASNGVFEVDKDSFYNNIIVARA DNKSKAIQD FVKAQVOTDEVEAE 264

QY 265 KKQFKDGYI 273
DB 295 QEKYKAVV 303

RESULT 9
US-09-252-991A-20200

/ Sequence 20200, Application US/09252991A

/ Patent No. 6551795

/ GENERAL INFORMATION:

/ APPLICANT: Marc J. Rubenfield et al.

/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

/ FILE REFERENCE: 107196.136

/ CURRENT APPLICATION NUMBER: US/09/252,991A

/ PRIOR FILING DATE: 1999-02-18

/ PRIOR APPLICATION NUMBER: US 60/074,788

/ PRIOR FILING DATE: 1998-02-18

/ NUMBER OF SEQ ID NOS: 33142

/ SEQ ID NO 20200

/ LENGTH: 271

/ TYPE: PRT

/ ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-20200

Query Match 35.8%; Score 494.5; DB 4; Length 271;
Best Local Similarity 41.4%; Pred. No. 1.1e-43;

Matches 110; Conservative 49; Mismatches 90; Indels 17; Gaps 4;

QY 9 ICALSGIALAGCSNOSNEPAAISTKAQTITKGVMAQGEQAVAE-VAGQVAKERNLTV 67

DB 19 LIALASVI-----SVSANAAEKLLVATPIPHAEILIELIKPTLAKGVLDQI 65

QY 68 ELVEFNDAFNPASVSKGELDANAMQHKPYLEKDSQEKGLNNLVIGNTFVYPLAGYSTX 127

DB 66 KV--FTDVQV 122

QY 128 IKTNELKQATIAVNDPSNIALALILEKQGLIKKQNTNLFTSLDIYENPKVIX 187

DB 123 YKSLLELDGATVAILPNSGNSGRALLLQKAGLKLKDPNNALATPDKIAENPKNLKFK 182

QY 168 EVDTSVARAIDVDVLAVNNNVAQVGLTASENGVEVEDSDSPYNNIIVARADNKSKA 247

DB 163 EIESLALPRVLDQVLDLINTNYALEALNPAKDALVLEDHSPYNNIIVARPNKSDA 242

QY 248 IQDFVAKYQTDVEAEAKKQFKDGYI 273

DB 243 LKGLSALTSPEVKAFIEKYYAGAVV 268

QY 248 IQDFVAKYQTDVEAEAKKQFKDGYI 273

DB 243 LKGLSALTSPEVKAFIEKYYAGAVV 268

QY 248 IQDFVAKYQTDVEAEAKKQFKDGYI 273

DB 243 LKGLSALTSPEVKAFIEKYYAGAVV 268

QY 248 IQDFVAKYQTDVEAEAKKQFKDGYI 273

DB 243 LKGLSALTSPEVKAFIEKYYAGAVV 268

QY 248 IQDFVAKYQTDVEAEAKKQFKDGYI 273

DB 243 LKGLSALTSPEVKAFIEKYYAGAVV 268

QY 248 IQDFVAKYQTDVEAEAKKQFKDGYI 273

DB 243 LKGLSALTSPEVKAFIEKYYAGAVV 268

QY 248 IQDFVAKYQTDVEAEAKKQFKDGYI 273

DB 243 LKGLSALTSPEVKAFIEKYYAGAVV 268

QY 248 IQDFVAKYQTDVEAEAKKQFKDGYI 273

DB 243 LKGLSALTSPEVKAFIEKYYAGAVV 268

QY 248 IQDFVAKYQTDVEAEAKKQFKDGYI 273

DB 243 LKGLSALTSPEVKAFIEKYYAGAVV 268

QY 248 IQDFVAKYQTDVEAEAKKQFKDGYI 273

DB 243 LKGLSALTSPEVKAFIEKYYAGAVV 268

QY 248 IQDFVAKYQTDVEAEAKKQFKDGYI 273

DB 243 LKGLSALTSPEVKAFIEKYYAGAVV 268

QY 248 IQDFVAKYQTDVEAEAKKQFKDGYI 273

DB 243 LKGLSALTSPEVKAFIEKYYAGAVV 268

QY 248 IQDFVAKYQTDVEAEAKKQFKDGYI 273

DB 243 LKGLSALTSPEVKAFIEKYYAGAVV 268

QY 248 IQDFVAKYQTDVEAEAKKQFKDGYI 273

DB 243 LKGLSALTSPEVKAFIEKYYAGAVV 268

QY 248 IQDFVAKYQTDVEAEAKKQFKDGYI 273

DB 243 LKGLSALTSPEVKAFIEKYYAGAVV 268

QY 248 IQDFVAKYQTDVEAEAKKQFKDGYI 273

DB 243 LKGLSALTSPEVKAFIEKYYAGAVV 268

QY 248 IQDFVAKYQTDVEAEAKKQFKDGYI 273

DB 243 LKGLSALTSPEVKAFIEKYYAGAVV 268

QY 248 IQDFVAKYQTDVEAEAKKQFKDGYI 273

DB 243 LKGLSALTSPEVKAFIEKYYAGAVV 268

QY 248 IQDFVAKYQTDVEAEAKKQFKDGYI 273

DB 243 LKGLSALTSPEVKAFIEKYYAGAVV 268

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071.035
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 277 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

US-09-071-035-30

Query Match 34.8%; Score 481; DB 4; Length 277;
Best Local Similarity 36.8%; Pred. No. 2.9e-42;

Matches 102; Conservative 62; Mismatches 101; Indels 12; Gaps 5;

QY 3 FKGINGICLALAGCSNOSNEPAAISTKAQTITKGVMAQGEQAVAE-VAGQVAKERNLTV 62

DB 4 FSLTGLIGLIVA-FTIAGCAGS-----VKDTKTEVTVGAVGTNDMESYKDLKKK- 56

QY 63 YNLTVELVEFNDAFNPASVSKGELDANAMQHKPYLEKDSQEKGLNNLVIGNTFVYPLA 122

DB 57 -NIDQVVEFTTYTQPNALAKETIDNAFQGITLDMYNNKHG- TKVVSIGNTVNAPLG 114

QY 123 GYSTKIKTNELKQATIAVNDPSNIALALILEKQGLIKKQNTNLFTSLDIYENPK 182

DB 115 IVANKIKDITKIKDGEIPIPNPTNGGRALLIQTAGLIKVDPAKQQLPTVSDITEKR 174

QY 183 KLVKEVDTSVARAIDVDVLAVNNNVAQVGLTASENGVEVE--DKDSPYNNIIVAR 239

DB 175 QKTIETLDTQYARLQVDYDASVINGAVDAGYTPDDDAITLBEVNEKAKYNNIIVAR 234

QY 240 ADNKSKAIDQFVAKYQTDVEAEAKKQFKDGYI 276

DB 235 EEDQENKLYQKVVEEYQCEETKVIATSKGANVEM 271

QY 240 ADNKSKAIDQFVAKYQTDVEAEAKKQFKDGYI 276

DB 235 EEDQENKLYQKVVEEYQCEETKVIATSKGANVEM 271

QY 240 ADNKSKAIDQFVAKYQTDVEAEAKKQFKDGYI 276

DB 235 EEDQENKLYQKVVEEYQCEETKVIATSKGANVEM 271

QY 240 ADNKSKAIDQFVAKYQTDVEAEAKKQFKDGYI 276

DB 235 EEDQENKLYQKVVEEYQCEETKVIATSKGANVEM 271

QY 240 ADNKSKAIDQFVAKYQTDVEAEAKKQFKDGYI 276

DB 235 EEDQENKLYQKVVEEYQCEETKVIATSKGANVEM 271

QY 240 ADNKSKAIDQFVAKYQTDVEAEAKKQFKDGYI 276

DB 235 EEDQENKLYQKVVEEYQCEETKVIATSKGANVEM 271

QY 240 ADNKSKAIDQFVAKYQTDVEAEAKKQFKDGYI 276

DB 235 EEDQENKLYQKVVEEYQCEETKVIATSKGANVEM 271

QY 240 ADNKSKAIDQFVAKYQTDVEAEAKKQFKDGYI 276

DB 235 EEDQENKLYQKVVEEYQCEETKVIATSKGANVEM 271

QY 240 ADNKSKAIDQFVAKYQTDVEAEAKKQFKDGYI 276

DB 235 EEDQENKLYQKVVEEYQCEETKVIATSKGANVEM 271

QY 240 ADNKSKAIDQFVAKYQTDVEAEAKKQFKDGYI 276

DB 235 EEDQENKLYQKVVEEYQCEETKVIATSKGANVEM 271

QY 240 ADNKSKAIDQFVAKYQTDVEAEAKKQFKDGYI 276

DB 235 EEDQENKLYQKVVEEYQCEETKVIATSKGANVEM 271

QY 240 ADNKSKAIDQFVAKYQTDVEAEAKKQFKDGYI 276

DB 235 EEDQENKLYQKVVEEYQCEETKVIATSKGANVEM 271

QY 240 ADNKSKAIDQFVAKYQTDVEAEAKKQFKDGYI 276

DB 235 EEDQENKLYQKVVEEYQCEETKVIATSKGANVEM 271

QY 240 ADNKSKAIDQFVAKYQTDVEAEAKKQFKDGYI 276

DB 235 EEDQENKLYQKVVEEYQCEETKVIATSKGANVEM 271

QY 240 ADNKSKAIDQFVAKYQTDVEAEAKKQFKDGYI 276

DB 235 EEDQENKLYQKVVEEYQCEETKVIATSKGANVEM 271

QY 240 ADNKSKAIDQFVAKYQTDVEAEAKKQFKDGYI 276

DB 235 EEDQENKLYQKVVEEYQCEETKVIATSKGANVEM 271

QY 240 ADNKSKAIDQFVAKYQTDVEAEAKKQFKDGYI 276

DB 235 EEDQENKLYQKVVEEYQCEETKVIATSKGANVEM 271

QY 240 ADNKSKAIDQFVAKYQTDVEAEAKKQFKDGYI 276

DB 235 EEDQENKLYQKVVEEYQCEETKVIATSKGANVEM 271

QY 240 ADNKSKAIDQFVAKYQTDVEAEAKKQFKDGYI 276

DB 235 EEDQENKLYQKVVEEYQCEETKVIATSKGANVEM 271

QY 240 ADNKSKAIDQFVAKYQTDVEAEAKKQFKDGYI 276

DB 235 EEDQENKLYQKVVEEYQCEETKVIATSKGANVEM 271

QY 240 ADNKSKAIDQFVAKYQTDVEAEAKKQFKDGYI 276

DB 235 EEDQENKLYQKVVEEYQCEETKVIATSKGANVEM 271

QY 240 ADNKSKAIDQFVAKYQTDVEAEAKKQFKDGYI 276

DB 235 EEDQENKLYQKVVEEYQCEETKVIATSKGANVEM 271

QY 240 ADNKSKAIDQFVAKYQTDVEAEAKKQFKDGYI 276

DB 235 EEDQENKLYQKVVEEYQCEETKVIATSKGANVEM 271

QY 240 ADNKSKAIDQFVAKYQTDVEAEAKKQFKDGYI 276

DB 235 EEDQENKLYQKVVEEYQCEETKVIATSKGANVEM 271

QY 240 ADNKSKAIDQFVAKYQTDVEAEAKKQFKDGYI 276

DB 235 EEDQENKLYQKVVEEYQCEETKVIATSKGANVEM 271

QY 240 ADNKSKAIDQFVAKYQTDVEAEAKKQFKDGYI 276

DB 235 EEDQENKLYQKVVEEYQCEETKVIATSKGANVEM 271

QY 240 ADNKSKAIDQFVAKYQTDVEAEAKKQFKDGYI 276

DB 235 EEDQENKLYQKVVEEYQCEETKVIATSKGANVEM 271

QY 240 ADNKSKAIDQFVAKYQTDVEAEAKKQFKDGYI 276

DB 235 EEDQENKLYQKVVEEYQCEETKVIATSKGANVEM 271

QY 240 ADNKSKAIDQFVAKYQTDVEAEAKKQFKDGYI 276

DB 235 EEDQENKLYQKVVEEYQCEETKVIATSKGANVEM 271

Tue Jun 22 11:04:04 2004

us-10-018-672-2.ra1

Page 5

QY 3 FGIINGICALAGLALAGCSNOSNEPAISKTAAQTIKVGWAGPEQVAEVAQVAKER 62
Db 5 FSKLIGLIGVLA-FTIACASGS-----VKDTKETVKGIVGKNDWESVMDRLKK- 57
QY 63 YNLVELVEFNDYAMPNSAVSKGELDANAMQHKPYLEKDSOEKLNLVYVGTFTVPLA 122
Db 58 -NIDLOLVEFTDYTOPNALAEKSIDINAFQHOIFLNNYKHEG-TKLVSIGNTVAPLG 115
QY 123 GYSTKIKTINELKOGATIAVNDPSNLABILILEKOGILTKDNTLFTSLDIVENPK 182
Db 116 IYANKLKITKOGGELIAPNDPTNGRILITQTAGLIVDPAKQOLPVSDITENKR 175
QY 183 KLVIKEVDTSVAARIDVDLAVNNNYAGVGLTASENGVE--DKDSPYNNIIVAR 239
Db 176 QKITELDATQOTARALQVDASVINSQVADAGYTPDKALFLSPVNEKAKPYNNIIVAR 235
QY 240 ADNNDSKAIQDPVYKQVDEVEAEAKKQFKDGVIKGM 276
Db 236 EEDQENKLYQKVEBYEQEETKVIATSKGANVPAM 272

RESULT 12
US-09-543-681A-4774
Sequence 4774, Application US/09543681A
Patent No. 6605709
GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543,681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 4774
LENGTH: 269
TYPE: PRT
ORGANISM: Proteus mirabilis
US-09-543-681A-4774

Query Match 34.3%; Score 475; DB 4; Length 269;
Best Local Similarity 35.8%; Pred. No. 1,2e+41;
Matches 96; Conservative 60; Mismatches 104; Indels 8; Gaps 4;
QY 12 LASGIALAGCSNOSNEPAISKTAAQTIKVGWAGPEQVAEVAQVAKERKNTLVEVE 71
Db 7 IITGLILASISLFS--VSCSQEDIDYKVAINTGPDQVIMDEVIRLAKTEGLINVEVIA 63
QY 72 FNDYAMPNSAVSKGELDANAMQHKPYLEKDSOEKLN-NUYVNTFTVPLAGSTIKT 130
Db 64 FNDYERPKALENRVDVNFQSTIPYDQ--SEIRAHNYKFKHTASKTIIFPLAAYSKISD 121
QY 131 INELKOGATIAVNDPSNLABILILEKOGILTKDNTLFTSLDIVENPKLIVKEVD 190
Db 122 ISELVEGVVAIPNEASVSKGELILLENHLSLKEGVGLPSVEDIINDNNALIFREVE 181
QY 191 TSVARALDD--VDLAVNNNYAGVGLTASENGVEVDKSPYNNIIVARADKDSKAI 248
Db 182 TPLMVEALDDPEVTMAVINNNFSQIGLATRDGLIMENKESPYANVVVTRIDNNDEKI 241
QY 249 QDFVKAQOTDEVEAEAKKQFKDGVIKGM 276
Db 242 KKLITVJHSRQVELKVGEMTKGDVAKAM 269

RESULT 13
US-09-071-035-32
Sequence 32, Application US/09071035
Patent No. 6448043
GENERAL INFORMATION:
APPLICANT: GILL H. Choi
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides

NUMBER OF SEQUENCES: 496
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,035
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 257 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-32

Query Match 33.4%; Score 462; DB 4; Length 257;
Best Local Similarity 38.0%; Pred. No. 2.6e+40;
Matches 95; Conservative 56; Mismatches 93; Indels 6; Gaps 3;
QY 30 AISKTAQTIKVGWAGPEQVAEVAQVAKERKNTLVEVEFNDYAMPNSAVSKGELDA 89
Db 5 SVMDTETEVKLGIVGKNDWESVMDRLKK--NIDLOLVEFTDYTOPNALAEKSIDL 62
QY 90 NAMQHKPYLEKDSOEKLNLVYVGTFTVPLAGSTIKTINELKOGATIAVNDPSNL 149
Db 63 NAFQHOIFLNNYKHEG-TKLVSIGNTVAPLGIYANKKDKITKIDGGEIAPNDPTNG 121
QY 150 ARALILILEKOGILTKDNTLFTSLDIVENPKLIVKEVDTSVAAARIDVDLAVNNN 209
Db 122 GRLLILDTAGLILKVDPAKQOLPVSDITENKROQLITELDQOTARALQVDASVINS 181
QY 210 YAGVGLTASENGVE--DKDSPYNNIIVARADKDSKAIQDFVKAQOTDEVEAEAK 266
Db 182 MAYDAGYTPDKALIFEPVNERAKPYNNIIVAREEOENKLYQKVEEYQOEBETKVIAE 241
QY 267 QFKDGVIKGM 276
Db 242 TSKGANVPAM 251

RESULT 14
US-09-107-532A-7234
Sequence 7234, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street

NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 5195:
SEQUENCE CHARACTERISTICS:
LENGTH: 280 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 5195:
US-08-956-171E-5195

Query Match 30.1%; Score 416; DB 4; Length 280;
Best Local Similarity 35.8%; Pred. No. 2,1e-35;
Matches 93; Conservative 50; Mismatches 101; Indels 16; Gaps 4;

QY 16 IALAGCNSQNEPRAISKTAQTIKVGMAGPEQAVAVAGVAKKYNLTVELVENDY 75
DB 13 VLLAACGAND-----KKVTIGVASNDTAMEKVK-ELAK-KDDIDVEIKHSY 60
QY 76 AMPNSAVSKGELDANAVQHKPYLEKDSQEKLNNTLVGNTFYVPLAGYSTIKITNELK 135
DB 61 NLPRKALNDGIDWNAFHAFADQYKKAHKGTISALSTTVLAPLGYSDKIKDYKVK 120
QY 136 DGATIAVNPDSNLRALLILEKQGLIKLKDNTNLFSTTIDYENPKVYIKYDVSVA 195
DB 121 DGAKVVIPIPNVSNQARALKLEAAGLIKTKDFELAGTVKDITSNPHLKITVDAQTA 180
QY 196 RAIDVVLAVNNVYAGVGLTASENGVEYEDKDS---PYVNIIVARADNKSKAIQDF 251
DB 181 RALSDVIAIYVNSVAKRAGCDPRNDPIFEKNSDAVKRYINIVANDDLNKTAKI 240
QY 252 VKAYQTDVEAEAKKQFQDG 271
DB 241 VELYHSEAKQALQEDVDG 260

RESULT 17
US-09-134-001C-4878
Sequence 4878, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 4878
LENGTH: 279
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4878

Query Match 27.0%; Score 373.5; DB 4; Length 279;
Best Local Similarity 33.9%; Pred. No. 6,1e-31;
Matches 92; Conservative 47; Mismatches 115; Indels 17; Gaps 7;

QY 9 ICALASGIALAGCNSQNEPRAISKTAQTIKVGVMAGPEQAVAVAGVAKK-KYLVTV 67
DB 17 ILVALITLALACG-----GKDKKITTIVASAPFAHIEIKKPKLIKKGIDYDKI 68
QY 68 ELVEFNDAVNPNSAVSKGELDANAVQHKPYLEKDSQEKLNNTLVGNTFYVPLAGYSTK 127

DB 69 KPI--NDVTPNKLLDKXGEIDANFPQHTPYNTBESKGY-KISBAGNVELEPNVAVSEK 125
QY 128 ITTLNELKDGTIANVPNDSNLRALLILEKQGLIKLKDNTNLFSTTIDYENPKVYI 186
DB 126 YSLDLPKGAIVVYVSNPABQGRFLKPFVDEGLIKLKGVKIEAKFDDITENKDIKF 185
QY 187 KEVDTSVAARAI---DVPDLAVNNVYAGVGLTASENGVEED-KDSPPYNIIVARADN 242
DB 166 NNKQSAEYIPIKYNQDADAVIYNSNVAIDQKLSPKXDSIALBESKDPYANLIVKKGH 245
QY 243 KDSKAIQDFVKAAYQTDVEAEAKKQFQDGYI 273
DB 246 KDDKNIKVLMETVLQSKETIDYIKDYKDAVV 276

RESULT 18
US-09-107-532A-6443
Sequence 6443, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSER: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD/ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Denke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 6443:
SEQUENCE CHARACTERISTICS:
LENGTH: 269 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...269
SEQUENCE DESCRIPTION: SEQ ID NO: 6443:
US-09-107-532A-6443

Query Match 26.5%; Score 366; DB 4; Length 269;
Best Local Similarity 35.5%; Pred. No. 3,5e-30;
Matches 89; Conservative 43; Mismatches 85; Indels 34; Gaps 7;

QY 42 GVMAGPEQAV-----AEVAGV-----AKKYNLTVELVENDY 74
DB 3 GRNRGQQLNRIPIVSCETTPPGALCSAPISPAEMGDHVNPLWTHGFDL--ETIGFDD 60

Tue Jun 22 11:04:04 2004

us-10-018-672-2.ra1

Page 9

```

QY 72 FNDYAMPNSAVSK-GELDDANNAQHKPYLEKSOEGLNNLVYGVNTVYVPLAGSTKTK 130
Db 68 FTTYNNOPNQTLLKXDDGDIQUNAFQHNFL-NMNMKKHNTKIIDIDTYIGPWRASNNKTKS 126
QY 131 LNELKOGATVAVENDBSNLRALITLEKOGILTKLKDNTLPSSTLLDIYENPKXLIYEND 190
Db 127 LKEXKGGDDQISLENDNSNGBRALQLLAQWKLITLKKGV-ASPTLRDITSKINKKFFELD 185
QY 191 TSYAARAIDDVDLVANNVNNYAGOVGLTASENGVFVE--DYDSEYVNIYVAAAD-NKSK 246
Db 186 AAGTAAASLNDVAAVAVVNNIDIAAANLKPS-NAAVEKITSQKPPVNFIAAKSKDKNNA 244
QY 247 AIQDPVKAYOT 257
Db 245 TYKKIVKAYOT 255

```

```

RESULT 22
US-08-961-083-32
Sequence 32, Application US/08961083
Patent No. 6159469
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,083
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 260 amino acids
TYPE: amino acid
STRADEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-961-083-32

```

Query Match	25.7%;	Score 356;	DB 3;	Length 260;
Best Local Similarity	35.9%;	Pred. No. 3.8e-29;		
Matches 90;	Conservative 50;	Mismatches 93;	Indels 18;	Gaps 8;

QY 36 AOTIKYGVN--AGPBOVAEVAQVAKKYNLTVELVENNDYAMPNSAVSKGELDANNO 93
Db 9 ATTIKIATVNRSGSEKEMDKIOELVK-KDGTLEFTEFTLEDSQPKKATPDGVDVDAIAQ 67
QY 94 HKPYLEKDSQOEKGLNNLVIAGNTFYVPLAGYS-----TXIKLNLKDGATTIAPENDS 147
Db 68 HYNFLNNMKENG-KDLVAIADTYISPIPLYSGLSANGANKYTKVEDIIPANGELAVENDAT 126
QY 148 NAFAPALILEKOGGLIKKDNNTLFTSLRIVEPPKCLVIKEVDTSVAAAIADVDLAVVN 207

Db	127	NESRALYLLQSGLLIDVSGTALATVANKENPKNLKITELDASQJARSISSDAAVVN	186
Qy	208	NNNAGOVGLTASENGVEVEDKD--SPVNIIVARAD--NKDSKAIQDPVKAYOTDEVE	261
Db	187	NTEVTEAKLDYKKS-LFKQADENSKQWNIIVAKKDQETSPPKDALIKVIAVAHTDDVK	245
Qy	262	AEAKQPFKDG V 272	
Db	246	KVIESSDGL 255	

```

1      RESULT 23
2      US-09-536-784-32
3      / Sequence 32, Application US/09536784
4      / Patent No. 6573082
5      / GENERAL INFORMATION:
6      / APPLICANT: Choi et. al.
7      / TITLE OF INVENTION: Streptococcus pneumoniae Antigenes and Vaccines
8      / NUMBER OF SEQUENCES: 452
9      / CORRESPONDENCE ADDRESS:
10     / ADDRESS: Human Genome Sciences, Inc.
11     / STREET: 9410 Key West Avenue
12     / CITY: Rockville
13     / STATE: Maryland
14     / COUNTRY: USA
15     / ZIP: 20850
16     / COMPUTER READABLE FORM:
17     / MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
18     / COMPUTER: HP Vectra 486/33
19     / OPERATING SYSTEM: MSDOS version 6.2
20     / SOFTWARE: ASCII Text
21     / CURRENT APPLICATION DATA:
22     / APPLICATION NUMBER: US/09/536,784
23     / FILING DATE: 30-Oct-1997
24     / CLASSIFICATION: <Unknown>
25     / PRIOR APPLICATION DATA:
26     / APPLICATION NUMBER: 08/961,083
27     / FILING DATE: OCT-30-1997
28     / ATTORNEY/AGENT INFORMATION:
29     / NAME: Michelle S. Marks
30     / REGISTRATION NUMBER: 41,971
31     / REFERENCE/DOCKET NUMBER: P8340P3
32     / TELECOMMUNICATION INFORMATION:
33     / TELEPHONE: (301) 309-8504
34     / TELEFAX: (301) 309-8512
35     / INFORMATION FOR SEQ ID NO: 32:
36     / SEQUENCE CHARACTERISTICS:
37     / LENGTH: 260 amino acids
38     / TYPE: amino acid
39     / STRANDEDNESS: single
40     / TOPOLOGY: linear
41     / MOLECULE TYPE: protein
42     / SEQUENCE DESCRIPTION: SEQ ID NO: 32:
43     US-09-536-784-32

```

Query Match	25.7%	Score 356	DB 4	Length 260
Best Local Similarity	35.9%	Pred. No. 3.8e-29		
Matches 90: Conservative	50:	Mismatches 93:	Indels 18:	Gaps 8:

Qy 36 QOTIVGM--AGPQAAVAEVAQVAKEXNLTVELVFNZYAMNSVSGELDANAMQ 93
Db 9 ATTIIATVNRSGSEEEKWDKI QELVVK-KDGIITLFTFETIDYSQPNKATAGEDVNAFQ 67
Qy 94 HKPYLEKDSOEKGLNLTIVGNTFVYPPLAGS-----TKITINELKDQATIAPNDPS 147
Db 68 HYNFLNNMKNKENG-KDLVLAIDPTTISPRLXSGANGSANKXTKYVEDIIPANGELIANPNAT 126
Qy 148 MLARALLLEKQGLIKLKDNNLTLPSTTLDIENPQKVLKIEVDTSVAAARIDVDLAVN 207
Db 127 NESRRLTYLQSGILKLDVSGTALATVANIENPNLKITELDASQTRARLSSVDAAVNN 186
Qy 208 NNVAQVGLTNSENGVEVEDKD---SPYVNITIVARAD---NKDQALIDFPYKAYITDEVE 261

```

Db      187 NTFTVATLADYDKKS-LFREQADENSKOMYNTIYAKKOWMISPPKADAIKVIYAAYHTDDK 245
QY      262 AEAKKQFQGV 272
        : : : : :
Db      246 -KVIESSDGL 255

RESULT 24
US-08-956-171E-5196
; Sequence 5196, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
;           Gil H. Choi
;           Patrick S. Dillon
;           Craig A. Rosen
;           Steven C. Barash
;           Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Human Genome Sciences, Inc.
;   STREET: 9410 Key West Avenue
;   CITY: Rockville
;   STATE: Maryland
;   COUNTRY: USA
;   ZIP: 20850
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
;   COMPUTER: HP Vectra 486/33
;   OPERATING SYSTEM: MSDOS version 6.2
;   SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/956,171E
;   FILING DATE: 20-Oct-1997
;   CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 60/009,861
;   FILING DATE: January 5, 1996
;   APPLICATION NUMBER: 08/781,986
;   FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
;   NAME: Mark J. Hyman
;   REGISTRATION NUMBER: 46,788
;   REFERENCE/DOCKET NUMBER: PB240P1
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (240) 314-1224
;   TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 5196:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 273 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   SEQUENCE DESCRIPTION: SEQ ID NO: 5196:
US-08-956-171E-5196

Query Match      25.0%; Score 346; DB 4; Length 273;
Best Local Similarity 31.0%; Pred. No. 4,6e-28;
Matches 85; Conservative 53; Mismatches 116; Indels 20; Gaps 8;

QY      9 ICALASGIALGSSNSQNEPAISKRAQTIVGVAVAGPEQAVAVAGOVAEK-YNLTV 67
Db      8 ILVLTAVVLAAGNGK-----SGSDDKKIVTSASPAFHADILEKAPLEKKYEID 62
QY      68 ELVEFNDYAMPNSAVSKGEIDANAMQHKFYLEKDSQEGKLNLVIVGNTFYFVPLAGVSTK 127
        : : : : :
Db      63 KTI--NDYTPNKLDKGEIDANYFQHTPYLMTKKDKGY-KIVSAGDVHLEPAVVSXK 119
        : : : : :
QY      128 IKTLNEKQCATIAYVNDPSNLARALLILEKGLIKLKNNTLV-----FSTLIDIVENPKK 183
        : : : : :

```

```

Db      120  IKSLKELPKGATTVVSNPNPAEQGRFLKEFPFDALILIKKGVKLEDAKFS---DITENKCD 176
Qy      184  LVYKEVDTSVAAAR1---DDVDLAVNNNYAGVGLTASENGVFED-KDSFYNNIYAR 239
Db      177  IKFNKNSAEPFLPKYONEDADAVIINSNPAIQKLNPKKDSIAVESAKDNPANLIAVK 236
Qy      240  ADNKSRAIODFVAVYQTDVEAEAKKQFMDGYI 273
Db      237  EGHODDKIKALIEVLVDSKDIODFINEKINGAVI 270

RESULT 25
US-09-328-352-6702
; Sequence 6702, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6702
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6702

Query Match      24.4%; Score 337.5; DB 4; Length 286;
Best Local Similarity 37.7%; Pred. No. 3,9e-27;
Matches 87; Conservative 40; Mismatches 89; Indels 15; Gaps 7;

Qy      33  KTAAGTIRYGV--MAGPEQAVAEVAGQVAKKYNLTVELVEFNDVPAMPNSAVSGEIDA 89
Db      41  QSGSNELVIGISPPFAKPLQAAADEA-----KKQGLNVLAVFSDMTNPNTILNHGIDIA 95
Qy      90  NAMQHKPYLEKDSQEKGLN-NLYVGNTPVYPLAGISTYKIKTLNELKDGATTAVPNPSN 148
Db      96  NFDQHPFLDNALKEGFGFKLKAFAVGAASHVGL--YSKRYKSLDELDPONARVVIENDEVN 153
Qy      149  LARALITLKEKGLIKLKDNMTNLFSTLTDIVENPKTLVKEVDTSVAAARIDVDLAVVNN 208
Db      154  QGRALLLQQAKITTKDSNNHLSALKDVSNPKLQFLYEVEGPTAAIIDVDLAFGYP 213
Qy      209  NYAGVGLTASENGVFEDKDS--PYNNIYARADNKD-SKAIQDFVAYQ 256
Db      214  HYV-RLAKTPADPNSALLLDNTNKRKAYALFVVRDDEYKGRKLKKFVEIYQ 263

RESULT 26
US-09-328-352-7028
; Sequence 7028, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7028
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7028

Query Match      24.2%; Score 335; DB 4; Length 304;
Best Local Similarity 32.0%; Pred. No. 7,9e-27;
Matches 89; Conservative 43; Mismatches 126; Indels 18; Gaps 5;

5 KINGCICALASGIALAGCSNQSNEPAISKTA-AQTIRGVNAGGEQAVAEVAGVANAKEY 63

```


Tue Jun 22 11:04:04 2004

us-10-018-672-2.rat

Page 11

Db 19 KTLISFLSVSVLLAACKCKQNEPQNGKDSKQLOTVIATSGPADWRIATLTPETKAA 78
Qy 64 NLTVLVEFENDYAMPNSAVSKGEJDANAMQHKPYL-----EKDSQEKGLNNLVIYGNTP 117
Db 79 GLKLEVKNFDTYVAMNTATANKVEDLNAFCGYAVLAFNNGND-----KIAPVATY 131
Qy 118 VYPLAGYSTKIKITNELKDGATIAVPNDPSNLARLLILEKQGLIKLKDNTNLFSTLD 176
Db 132 LEPMGITSSKTKKVDPEPQGSIAIPNDANEAVALLLLSQAGLITKADPDPKGPSPD 191
Qy 177 IVENPKLVIKEVDTSVAARAIDVDLAVNNNNAQGVGLTASENGVEFEDKSP---YV 233
Db 192 ITDSSKIKIDIKPIGMATRAVRKDEVDALVIGNTLAMEGLNVLKDSIYEPVQSTKLV 251
Qy 234 NIIVARADNKDSKAIQDFVKAQYQDEVEAEAKKQF 268
Db 252 NILATASRKDDPVLQKVGQLYHTEAVKXVECHF 286

RESULT 27

US-09-328-352-5549
; Sequence 5549, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Brelton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328.352
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5549
; LENGTH: 294
; TYPE: PRF
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5549

Query Match 23.5%; Score 325.5; DB 4; Length 294;
Best Local Similarity 32.9%; Pred. No. 7.4e-26;
Matches 91; Conservative 48; Mismatches 119; Indels 19; Gaps 8;

Qy 5 KINGICLAGGIALAGCSNOSPRAISKTA-----OTIKVWAGPEQAVAE-VAG 56
Db 6 KLFVFLRSASVLTTLTACNKKQ---PAQENTNAKDKTESVRTIKL-VSTGSDTDVWKXVAT 61
Qy 57 QVAREKYNLTVLVEFNDYAMPNSAVSKGELDANAMQHKPYLER-DSQEKGLNNLVIYGN 115
Db 62 LPETRAQIGIKLEVTNLTDYVLTNSVASEODVNAFQSPNYLAIYASNSKA--KVAAVAT 119
Qy 116 TVVYPLAGYSTKIKITNELKDGATIAVPNDPSNLARLLILEKQGLIKLKDNTN-LFSTT 174
Db 120 TYLEPMGITANKYKTVDEFPQGSATIAIPNDTANEARALLTLQSKAKIKLKPFDPVYGIV 179
Qy 175 LDIYENPKLVIKEVDTSVAARAIDVDLAVNNNNAQGVGLTASENGVEFEDKSP--- 231
Db 180 NDIENENPKNLQKPIQMTTAVRVKNDVDALVIGNTLAMEGLNVMKDALFREPIDOSTKL 239
Qy 232 YNIIIVARADNKDSKAIQDFVKAQYQDEVEAEAKKQF 268
Db 240 YVNLIGVAEANKNDPIYTKLGEIYHLPKXQKFNVEKF 276

RESULT 28
US-09-071-035-22
; Sequence 22, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,035
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 272 amino acids
TYPE: amino acid
STANDARDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-22

Query Match

23.4%; Score 324; DB 4; Length 272;
Best Local Similarity 30.3%; Pred. No. 9.4e-26;
Matches 81; Conservative 56; Mismatches 104; Indels 26; Gaps 8;

Qy 16 ILAGCSNOSPRAISKTAQTIKVGWAGPEQAVAEVAGQVAK-EKYNLTVLVEFVN 73
Db 17 LVIGAGCNKKKSDS-----VLKVGASPVPH---ABILEHVKKLLEKGVKLEVTYTT 65
Qy 74 DYAMPNSAVSKGELDANAMQHKPYLERKDSQEKGLNNLVIYGNTPFYPLAGYSTKIKITNE 133
Db 66 DVLPRKALIESGDIDANFYOHVPFENEAVKENDY-DFVWAGIHLPEPLGLSKKYSLOE 124
Qy 134 LKDGATIAVPNDPSNLARLLILEKQGLIKLKDNTNLFSTLD-DIYENPKLVIK-EYDT 191
Db 125 IPDGSTIYSSSVSDMPRTLLEDAGLITLKEGVDRRTATPDDIDDKTKLKRFHESDP 184
Qy 192 SVAARAIDVDLIA--VNNNNAQGVGLTASENGVEFEDKSPYNNIIVARADNKDSKAIQ 249
Db 185 AIMTLYDNEEGAAVLINSNFAVDQGLNPKKDALALEKSSPYANNIIVAREDENNERVYK 244
Qy 250 DFVKAQYQDEVEAEAKKQFQGVIKGW 276
Db 245 KLVKVLRSKEVO-----DWITKKW 263

RESULT 29

US-09-252-991A-25420
; Sequence 25420, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,035
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8512
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 251 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-24

Query Match 23.0%; Score 318; DB 4; Length 251;
Best Local Similarity 30.5%; Pred. No. 3,5e-25;
Matches 80; Conservative 55; Mismatches 101; Indels 26; Gaps 8;

QY 21 CSNOSNEPAISKTAQTIKTVAGVAGPEQAVAEVAGVAK--EKYNLTVELVEFNDYAMP 78
DB 1 CGNKSDSDS-----VLKVGASVPRH---AEIIEHKPLLEKGVLTETTTDYLP 49
QY 79 NSAYSKGLDANAMOHKRYLEKDSQEGKLNLVVGNFTVFPPLAGYSTKIKTLNELDGA 138
DB 50 NKALSGGIDANFYQHVPFFNEAVKENDY-DFVNAGAHLEPVLGYSKYSGLQEIPIDG 108
QY 139 TIAVNDPSNLARALLLEKQGLIKLKDNNTLFTTL-DIVENPKKIVIK-EVDTSVAR 196
DB 109 TIYSSSVSDMPRVLTILEDAGLITLKGVTRTATDDIDKNTKKLKNHESDPALMTT 168
QY 197 AIDVDLA--VNNNYAGQVGLTASENGVEFEDKSPYNIIVARADNKSKAIQDPVKA 254
DB 169 LYDNEEGALVILNSNFAVDQGLNPKDAIALEKSSPYANIIVARKDENNEVKLVKV 228
QY 255 YQDVEVEAEAKQKQFQVIGK 276
DB 229 LRSKEVQ-----DWITTKW 242

RESULT 33

US-09-071-035-10
Sequence 10, Application US/09071035
Patent No. 6448043

GENERAL INFORMATION:

APPLICANT: Gil H. Choi

TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides

NUMBER OF SEQUENCES: 496

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/071,035

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 272 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-10

Query Match 22.8%; Score 315.5; DB 4; Length 272;
Best Local Similarity 31.9%; Pred. No. 7,4e-25;
Matches 84; Conservative 52; Mismatches 116; Indels 11; Gaps 7;

QY 11 ALASGIALAGCSNOSNEPAISKTAQTIKTVAGVAGPEQAVAEVAGVAKKYNLTVELV 70
DB 10 AVIATVILAACG--GNKQA--DQKEDKEITVAVOLSSKDIETIAKKA-EKKGYKINIM 64
QY 71 EFNQVAMPNSAVSKGLDANAMOHKRYLEKDSQEGKLNLVVGNFTVFPPLAGYSTKIKT 130
DB 65 EYSDNVAINDAVOHDEADANFPAQHOFEMEFNKKA-DLVAOPIYYAGGFYSKEYD 123
QY 131 LNELKQATIAVNDPSNLARALLLEKQGLIKLKDNNTLFTTLDIVENPKKIVIK 190
DB 124 AKDLPENAKVGRSPDPTNEGRLALILNANGVILKLEGVGFNGTVADVNPKNITPESID 183
QY 191 TSVARAIIDVDLAVI--NNNTYAGQVGLTASENGVEFEDKSP--YNIIVARADNKSK 246
DB 184 LNLIAVAYDEKDIAMFCYPALTEPAGLT-TYDAITLLEKSKSKYVALQVTRKGEKSE 242
QY 247 AIQDFYKAYQDVEVEAEAKQK 269
DB 243 KIKVLKEAMTTKEVAVIYIKNSK 265

RESULT 34

US-09-071-035-12
Sequence 12, Application US/09071035
Patent No. 6448043

GENERAL INFORMATION:

APPLICANT: Gil H. Choi

TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides

NUMBER OF SEQUENCES: 496

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/071,035

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: A. Anders Brookes

REGISTRATION NUMBER: 36,373

REFERENCE/DOCKET NUMBER: PB369P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 253 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-12

Query Match 22.3%; Score 308.5; DB 4; Length 253;
Best Local Similarity 33.2%; Pred. No. 3,6e-24;
Matches 78; Conservative 46; Mismatches 104; Indels 7; Gaps 5;

39 IKVGVAPGBOVAEVAAGVAKKLNLTVELVEFNDYAMPNSAVSKGELDANAMQHKPYL 98
15 ITVAVOLESSKDIEIAKKA-EKKGYKINMEVSDNVAVDADQCHDEADANFAQHOPFM 73
99 EKDSQKGLNNLVVGNFTVYPLAGYSTIKITLNEKOGATIAVNDPSNLARLLILEK 158
74 EWFNKKKA-DLVAQPIYFAGGFYSKEDADLPENAKVGIIPSDPTNKGRLATILNA 132
159 OGILKNDNTNFTTLIVENPKKLVKEVDTSVAARIDVDLAVV--NNYAGQVGL 216
133 NGVILKKGVGNGTVADVNPKNITFESIDLNTAKAYDEKDIAYFCYPALEPAGL 192
217 TASENGVEVEDKSP--YVNIIVARADNKSRAIQDFVAYQTBVEAKKQK 269
199 T-TKDAILEDKASKHYALQVTRKGEKSEKIKVLEAMTKFEVALEYIKNSK 246

RESULT 35
US-09-134-000C-5006
Sequence 5006, Application US/09134000C
Patent No. 6617156

GENERAL INFORMATION:
APPLICANT: Lynn Doucet-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
FILE REFERENCE: 032796-032
CURRENT FILING DATE: 1998-08-13
CURRENT APPLICATION NUMBER: US/09/134,000C
PRIOR FILING DATE: 1997-08-15
PRIOR APPLICATION NUMBER: US 60/055,778
NUMBER OF SEQ ID NOS: 6812
SOFTWARE: Patent version 3.1
SEQ ID NO 5006
LENGTH: 278
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-134-000C-5006

Query Match 21.8%; Score 301.5; DB 4; Length 278;
Best Local Similarity 31.5%; Pred. No. 2,3e-23;
Matches 78; Conservative 52; Mismatches 107; Indels 11; Gaps 7;

11 AASGALAGCSNQSNEPAISKTAQTIKVGVMAGPEQVAEVAAGVAKKXNLTVELV 70
14 AVIATVILACG--GKKA--DQKEDKETTAVOLESSKDIEIAKKA-EKKGYKINIM 68
71 EFNDAVMPNSAVSKGELDANAMQHKPYLEKDSQKGLNNLVVGNFTVYPLAGYSTIKIT 130
69 EVDNVAVDADQCHDEADANFAQHOPFMENKKA-DLVAQPIYFAGGFYSKEDYOD 127
131 LNEIKGATIAVNDPSNLARLLILEKOGILKNDNTNFTTLIVENPKKLVKEVD 190
128 AKDLPENAKVGIIPSDPTNKGRLATILNANGVILKKGVGNGTVADVNPKNITFESID 187
191 TSAARIDVDLAVV--NNYAGQVGLTASENGVEVEDKSP--YVNIIVARADNKS 246

DB 188 LNLAKVDEKDIAMVFCYPALEPAGLT-TKDAILEDKASKHYALQVTRKGEKSE 246
QY 247 AIQDFVKA 254
DB 247 KIDGFKRS 254

RESULT 36
US-09-198-452A-291
Sequence 291, Application US/09198452A
Patent No. 6559294

GENERAL INFORMATION:
APPLICANT: Griflais, R.
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
FILE REFERENCE: 9710-003-999
CURRENT FILING DATE: 1998-11-24
CURRENT APPLICATION NUMBER: US/09/198,452A
NUMBER OF SEQ ID NOS: 6849
SEQ ID NO 291
LENGTH: 275
TYPE: PRT
ORGANISM: Chlamydia pneumoniae
US-09-198-452A-291

Query Match 18.6%; Score 257.5; DB 4; Length 275;
Best Local Similarity 32.2%; Pred. No. 9,5e-19;
Matches 77; Conservative 49; Mismatches 98; Indels 15; Gaps 9;

33 KINAQTIKVGWAGPEQVAEVAAGVAKKXNLTVEL--VEFNDYAMPNSAVSKGELDAN 90
23 KEDAQN-KIRIVASPTPH-AELLESQDEAKDGLIKLIPVDDIRIRRLLDQVADN 80
91 AMQHKPYLEKDSOE--KGLNNLVVGNFTVYPLAGYSTIKITLNEKOG--GATIAVND 145
81 YFQHOAFLEDECEERYDCKG--ELVVIYAKVHLBPQATYSKHSLSRLKSQKLTITAI 138
146 PSNLARLLILEKOGILKNDNTNFTTLIV--ENPKLVIXKVDNLSVAARIDVDVL 203
139 RTNAGRALHLECCGILVCKGPNLMTAKDVCGKEN-RSTVILEVSAPLIVGSLPDDA 197
204 AVANNYAGQVGLTASENGVEVEDKSP--YVNIIVARADNKSRAIQDFVAYQTBVE 261
198 AVIPGNFALANLSPKDKSLCEDLSVSKYTNLTVIRSDVSPKNIKRLQKLFQSPSVQ 256

RESULT 37
US-09-489-039A-10023
Sequence 10023, Application US/09489039A
Patent No. 6610836

GENERAL INFORMATION:
APPLICANT: Gary Breton et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709,2004001
CURRENT FILING DATE: 2000-01-27
CURRENT APPLICATION NUMBER: US/09/489,039A
PRIOR FILING DATE: 1999-01-29
PRIOR APPLICATION NUMBER: US 60/117,747
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 10023
LENGTH: 306
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10023

Query Match 15.1%; Score 209; DB 4; Length 306;
Best Local Similarity 28.7%; Pred. No. 1,4e-13;
Matches 71; Conservative 39; Mismatches 109; Indels 28; Gaps 8;

38 TIKV--GWAAGPEQVAEVAAGVAKKXNLTVELVEFNDYAMPNSAVSKGELDANAMQ 94

Db 61 TLKIHPEPAMAG-EORIIIEYINCHIAPDYLGLLEAVGVODPVQADRAVAEGVAGTIYQH 119
QY 95 KEYLEK--DSOEKGNINLVYVNTGVYPLAGYSTKIKTLNELKQATIAVNDPSNLARA 152
Db 120 QMWLQVVDANGFALSTVYV---FQMAFGIYSDRYSVQALPNCATIVPEDDGNQGA 176
QY 153 LILEKQGLIKLKDNTLNFSTTL-DIVENPKQVYKEVDTSVAAARAIIDVDLAVNNNYA 211
Db 177 LMLVORIGLISIDPAVEPRITAKNIYGNPHQFVKEHDLTMPRALNSVDAAI---GYV 233
QY 212 GO--VGLASENGVEDEKDSFYNNIIVARADNDSKAIQDFVAYQDVEAEAKQFK 269
Db 234 SDFADKVPREKGLIFPPAPRTFASQVITGT-----PYLSQENIVLTKQAFS 280
QY 270 DGVIKGW 276
Db 281 DRIQTW 287
RESULT 38
US-09-134-000C-3687, Application US/09134000C
; Sequence 3687, Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134, 000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3687
; LENGTH: 91
; TYPE: PRN
; ORGANISM: Enterococcus faecalis
US-09-134-000C-3687

Query Match 7.8%; Score 108; DB 4; Length 91;
Best Local Similarity 25.8%; Pred. No. 0.0085;
Matches 23; Conservative 25; Mismatches 31; Indels 10; Gaps 2;

QY 190 DTSVAAARAIIDVDLA--VNNNNAGVGLTASENGVEDEKDSFYNNIIVARADNDSKA 247
Db 2 DAAVATTLIDNEBGAVALINSFVADQGLNPKQALALEKSSPYANIIVAKEDENNNEN 61
QY 248 IQDFVAYQDVEAEAKQFKDGVIKGW 276
Db 62 VKKLIVKLSKEVQ-----DWITKKM 82
RESULT 39
US-09-328-352-6764
; Sequence 6764, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328, 352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6764
; LENGTH: 347
; TYPE: PRN
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6764

Query Match 7.7%; Score 106; DB 4; Length 347;
Best Local Similarity 26.0%; Pred. No. 0.012;

Matches 74; Conservative 33; Mismatches 118; Indels 60; Gaps 14;
QY 11 ALASGIALAGCSQNSNEPAA--ISKTPAQT--IKGVMAQPEQAVAEVAGQVAKER---- 62
Db 17 SILSAVVLGGCDTATKVPPEAKQDAQAANTKPIITIGYSDP-----GWVAMQVAILKGMWLK 72
QY 63 -YNIIVELVEFNPDYAMPNSAVSGELDANAMOKPYLEKDSOEKGLNINLVY-----GN 115
Db 73 EAGLNIPEKMF-DYSAISLSFSAHQDPAVLVNGDVLVAS--GGTQGMNLTATDYSAGN 129
QY 116 TFPVPLAGSTIKIKTLNELKQATIAVNDPSNLARAILLEKQGLIKLKDNTLNFSTTL 175
Db 130 DVIIIAKEG---INTIIDLK-GKSGIV-----EKLAV-----DHLLATLAL 165
QY 176 ---DIVENPKQVYKEVDTSVAAARAIIDVDLAVNNNNAGVGLTASENGVEDEKDSF- 231
Db 166 TDHNIKSNVEYKLVNSATNQDPQVFNSPDISAIAWQPVAGQALKAVAGSKIIITYSKDKPG 225
QY 232 -----YNNIIVARADNDSKAI---QDFVAYQDVEAEAKK 266
Db 226 LIYDTLTVNMSHLTAHQEEMKILIQWDKTVKYNPDATADAVK 270

RESULT 40
US-09-134-000C-5518
; Sequence 5518, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134, 000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5518
; LENGTH: 794
; TYPE: PRN
; ORGANISM: Enterococcus faecalis
US-09-134-000C-5518

Query Match 7.6%; Score 105; DB 4; Length 794;
Best Local Similarity 22.5%; Pred. No. 0.059;
Matches 64; Conservative 51; Mismatches 131; Indels 38; Gaps 10;

QY 1 MNFGKINGICALASGIALAGCSQNSNEPAAISKTAQITIKGVMAQPEQAVAEVAGQVAK 60
Db 387 VNNNILLRIBINGKRAQVKSNDAPDQGSNMAWMAAQF-GVGVSAFPKQ-LENGKNNFLK 444
QY 61 EKYNLVLEVEFNPDYAMPNSAV---SKGELDANAMOKPYLEKDSOEKGLNINLVY----- 111
Db 445 DNYV---LLBGNVPEKETDVTLYDSNNNTNINAKLGLGPDVKNQGLIARSDIYGTMK 500
QY 112 TFGNTPVPLAGSTIKIKTLNELKQATIAVNDPSNLARAILLEKQGLIKLKDNTLNFSTTL 169
Db 501 LANNNAFY-----TKLPTGNFIPIQDQAVYDNEBNE-----LTISGILIRIKSDSTN 549
QY 170 LESTLIDIVENPKQVYKEVDTSVAAARAIIDVDLAVNNNNAGVGLTASENGVEDEKDSF 229
Db 550 LLAPGIAVSDALSTIMIAKNTSDIVKQASNTWMTNNE---TLDASAKENLSTLYGAN 606
QY 230 SPYNNIIVARADNDSKAIQDFVAYQDVEAEAKQFKDGV 273
Db 607 EIPSSIMYFNDPKSKILDYLDAYNKG-----KDKKQI 643

Search completed: June 16, 2004, 11:09:17
Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 16, 2004, 11:09:22 ; Search time 17 Seconds

(without alignments)
845.374 Million cell updates/sec

Title: US-10-018-672-2

Perfect score: 276
Sequence: 1 NMFGKINGICALASGIALAG.....TDEVEAKKPKDGVIKGM 276

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 141681 seqs, 52070155 residues

Word size : 0

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	34	12.3	277 1	PLPB_PASHA
2	19	6.9	273 1	METQ_HAEN
3	12	4.3	263 1	PLPB_PASHA
4	12	4.3	276 1	METQ_PASHA
5	11	4.0	271 1	METQ_ECO57
6	11	4.0	271 1	METQ_SALTI
7	11	4.0	271 1	METQ_SALTI
8	9	3.3	125 1	RPA3_YEAST
9	9	3.3	269 1	METQ_VIBCH
10	9	3.3	391 1	MUCB_HUMAN
11	8	2.9	74 1	YD97_CAMJE
12	8	2.9	271 1	METQ_ECOLI
13	8	2.9	837 1	SMAG_MOUSE
14	8	2.9	838 1	SMAG_HUMAN
15	8	2.9	1220 1	LBX_MOUSE
16	7	2.5	103 1	IHEB_XANAC
17	7	2.5	103 1	IHEB_XANCP
18	7	2.5	104 1	IHEB_XYLFPA
19	7	2.5	104 1	IHEB_XYLFPT
20	7	2.5	141 1	PSPI_PSEIC
21	7	2.5	193 1	COAT_AFLSA
22	7	2.5	233 1	MRXB_KLEPN
23	7	2.5	240 1	CLCR_RAT
24	7	2.5	268 1	CLCR_RAT
25	7	2.5	268 1	SAPF_ECOLI
26	7	2.5	268 1	SAPF_ECOLI
27	7	2.5	269 1	SAPF_SALTI
28	7	2.5	271 1	METQ_YEAP
29	7	2.5	276 1	PLPB_PASHA
30	7	2.5	291 1	NAK2_ECOL6
31	7	2.5	297 1	ADT1_BOVIN
32	7	2.5	297 1	ADT1_BOVIN
33	7	2.5	298 1	ADT2_HUMAN
34	7	2.5	298 1	ADT2_MOUSE
35	7	2.5	298 1	ADT2_RAT
36	7	2.5	298 1	ADT3_BOVIN
37	7	2.5	298 1	ADT3_HUMAN
38	7	2.5	311 1	HEN3_NEIMA
39	7	2.5	311 1	HEN3_NEIMA
40	7	2.5	312 1	TFB1_THEAC
41	7	2.5	312 1	TFB1_THEAC
42	7	2.5	313 1	MAPL_CABEL
43	7	2.5	314 1	ARCC_LACSEL
44	7	2.5	341 1	FAH1_DEIRA
45	7	2.5	347 1	HPT_ATRGE
46	7	2.5	347 1	HPT_PABIT
47	7	2.5	348 1	FBPC_ECOLI
48	7	2.5	350 1	FORF_PSEAE
49	7	2.5	412 1	YG08_SYNY3
50	7	2.5	419 1	SVS_MYCTU
51	7	2.5	454 1	GATF_MERTH
52	7	2.5	457 1	DBDR_XENLA
53	7	2.5	471 1	YB08_YEAST
54	7	2.5	473 1	MBIN_THEMA
55	7	2.5	475 1	DBDR_RAT
56	7	2.5	477 1	DBDR_HUMAN
57	7	2.5	491 1	GALT_STEMU
58	7	2.5	538 1	PYRG_AERPE
59	7	2.5	550 1	SYR_MYCSM
60	7	2.5	575 1	TRAM_HAEN
61	7	2.5	598 1	VE1_HPV65
62	7	2.5	599 1	VE1_HPV65
63	7	2.5	692 1	GLND_CORGL
64	7	2.5	783 1	XPB_MOUSE
65	7	2.5	852 1	GLND_NEIMA
66	7	2.5	852 1	GLND_NEIMA
67	7	2.5	860 1	ATSE_HUMAN
68	7	2.5	889 1	CHSC_ASPEU
69	7	2.5	1239 1	DPG1_HUMAN
70	7	2.5	1972 1	BA2B_HUMAN
71	7	2.5	2130 1	BA2B_CHICK
72	7	2.2	61 1	AKH2_SCHNI
73	7	2.2	62 1	RJ32_METEX
74	7	2.2	63 1	DN72_STUSH
75	7	2.2	64 1	DN72_SULAC
76	7	2.2	65 1	DN71_SULAC
77	7	2.2	68 1	RR18_CVACA
78	7	2.2	69 1	GVPA_ANCAO
79	7	2.2	71 1	GVPA_PSEAN
80	7	2.2	72 1	YVDD_VACCV
81	7	2.2	73 1	Y43_BPR18
82	7	2.2	73 1	Y4GB_RHISM
83	7	2.2	79 1	YKGB_VACCOC
84	7	2.2	84 1	RS20_LISMO
85	7	2.2	88 1	FYV3_RAT
86	7	2.2	90 1	YVBD_ECOLI
87	7	2.2	92 1	RR19_GUTTH
88	7	2.2	92 1	RS18_DEIRA
89	7	2.2	95 1	VINI_BPT4
90	7	2.2	97 1	PY_DICLA
91	7	2.2	97 1	YOAF_BACSU
92	7	2.2	98 1	HUPE_RHITU
93	7	2.2	98 1	POOD_PRESM
94	7	2.2	100 1	URE2_MYCEB
95	7	2.2	101 1	URE2_MYCEB
96	7	2.2	102 1	YMSA_RHIME
97	7	2.2	106 1	Y04K_BPT4
98	7	2.2	107 1	KEE2_ECOLI
99	7	2.2	109 1	Y030_MYCTU
100	7	2.2	111 1	B2MG_RALRG
101	7	2.2	116 1	RT11_GHOCR
102	7	2.2	117 1	PAND_BRAJA
103	7	2.2	117 1	RL17_AQUAP
104	7	2.2	118 1	RL17_AQUAP
105	7	2.2	121 1	YF39_PYROH
106	7	2.2	122 1	RL18_SYNY3
107	7	2.2	122 1	RL18_SYNY3
108	7	2.2	122 1	RL18_SYNY3
109	7	2.2	122 1	RL18_SYNY3
110	7	2.2	122 1	RL18_SYNY3
111	7	2.2	122 1	RL18_SYNY3
112	7	2.2	122 1	RL18_SYNY3
113	7	2.2	122 1	RL18_SYNY3
114	7	2.2	122 1	RL18_SYNY3
115	7	2.2	122 1	RL18_SYNY3
116	7	2.2	122 1	RL18_SYNY3
117	7	2.2	122 1	RL18_SYNY3
118	7	2.2	122 1	RL18_SYNY3
119	7	2.2	122 1	RL18_SYNY3
120	7	2.2	122 1	RL18_SYNY3
121	7	2.2	122 1	RL18_SYNY3
122	7	2.2	122 1	RL18_SYNY3
123	7	2.2	122 1	RL18_SYNY3
124	7	2.2	122 1	RL18_SYNY3
125	7	2.2	122 1	RL18_SYNY3
126	7	2.2	122 1	RL18_SYNY3
127	7	2.2	122 1	RL18_SYNY3
128	7	2.2	122 1	RL18_SYNY3
129	7	2.2	122 1	RL18_SYNY3
130	7	2.2	122 1	RL18_SYNY3
131	7	2.2	122 1	RL18_SYNY3
132	7	2.2	122 1	RL18_SYNY3
133	7	2.2	122 1	RL18_SYNY3
134	7	2.2	122 1	RL18_SYNY3
135	7	2.2	122 1	RL18_SYNY3
136	7	2.2	122 1	RL18_SYNY3
137	7	2.2	122 1	RL18_SYNY3
138	7	2.2	122 1	RL18_SYNY3
139	7	2.2	122 1	RL18_SYNY3
140	7	2.2	122 1	RL18_SYNY3
141	7	2.2	122 1	RL18_SYNY3
142	7	2.2	122 1	RL18_SYNY3
143	7	2.2	122 1	RL18_SYNY3
144	7	2.2	122 1	RL18_SYNY3
145	7	2.2	122 1	RL18_SYNY3
146	7	2.2	122 1	RL18_SYNY3
147	7	2.2	122 1	RL18_SYNY3
148	7	2.2	122 1	RL18_SYNY3
149	7	2.2	122 1	RL18_SYNY3
150	7	2.2	122 1	RL18_SYNY3
151	7	2.2	122 1	RL18_SYNY3
152	7	2.2	122 1	RL18_SYNY3
153	7	2.2	122 1	RL18_SYNY3
154	7	2.2	122 1	RL18_SYNY3
155	7	2.2	122 1	RL18_SYNY3
156	7	2.2	122 1	RL18_SYNY3
157	7	2.2	122 1	RL18_SYNY3
158	7	2.2	122 1	RL18_SYNY3
159	7	2.2	122 1	RL18_SYNY3
160	7	2.2	122 1	RL18_SYNY3
161	7	2.2	122 1	RL18_SYNY3
162	7	2.2	122 1	RL18_SYNY3
163	7	2.2	122 1	RL18_SYNY3
164	7	2.2	122 1	RL18_SYNY3
165	7	2.2	122 1	RL18_SYNY3
166	7	2.2	122 1	RL18_SYNY3
167	7	2.2	122 1	RL18_SYNY3
168	7	2.2	122 1	RL18_SYNY3
169	7	2.2	122 1	RL18_SYNY3
170	7	2.2	122 1	RL18_SYNY3
171	7	2.2	122 1	RL18_SYNY3
172	7	2.2	122 1	RL18_SYNY3
173	7	2.2	122 1	RL18_SYNY3
174	7	2.2	122 1	RL18_SYNY3
175	7	2.2	122 1	RL18_SYNY3
176	7	2.2	122 1	RL18_SYNY3
177	7	2.2	122 1	RL18_SYNY3
178	7	2.2	122 1	RL18_SYNY3
179	7	2.2	122 1	RL18_SYNY3
180	7	2.2	122 1	RL18_SYNY3
181	7	2.2	122 1	RL18_SYNY3
182	7	2.2	122 1	RL18_SYNY3
183	7	2.2	122 1	RL18_SYNY3
184	7	2.2	122 1	RL18_SYNY3
185	7	2.2	122 1	RL18_SYNY3
186	7	2.2	122 1	RL18_SYNY3
187	7	2.2	122 1	RL18_SYNY3
188	7	2.2	122 1	RL18_SYNY3
189	7	2.2	122 1	RL18_SYNY3
190	7	2.2	122 1	RL18_SYNY3
191	7	2.2	122 1	RL18_SYNY3
192	7	2.2	122 1	RL18_SYNY3
193	7	2.2	122 1	RL18_SYNY3
194	7	2.2	122 1	RL18_SYNY3
195	7	2.2	122 1	RL18_SYNY3
196	7	2.2	122 1	RL18_SYNY3
197	7	2.2	122 1	RL18_SYNY3
198	7	2.2	122 1	RL18_SYNY3
199	7	2.2	122 1	RL18_SYNY3
200	7	2.2	122 1	RL18_SYNY3
201	7	2.2	122 1	RL18_SYNY3
202	7	2.2	122 1	RL18_SYNY3
203	7	2.2	122 1	RL18_SYNY3
204	7	2.2	122 1	RL18_SYNY3
205	7	2.2	122 1	RL18_SYNY3
206	7	2.2	122 1	RL18_SYNY3
207	7	2.2	122 1	RL18_SYNY3
208	7	2.2	122 1	RL18_SYNY3
209	7	2.2	122 1	RL18_SYNY3
210	7	2.2	122 1	RL18_SYNY3
211	7	2.2	122 1	RL18_SYNY3
212	7	2.2	122 1	RL18_SYNY3
213	7	2.2	122 1	RL18_SYNY3
214	7	2.2	122 1	RL18_SYNY3
215	7	2.2	122 1	RL18_SYNY3
216	7	2.2	122 1	RL18_SYNY3
217	7	2.2	122 1	RL18_SYNY3
218	7	2.2	122 1	RL18_SYNY3
219	7	2.2	122 1	RL18_SYNY3
220	7	2.2	122 1	RL18_SYNY3
221	7	2.2	122 1	RL18_SYNY3
222	7	2.2	122 1	RL18_SYNY3
223	7	2.2	122 1	RL18_SYNY3
224	7	2.2	122 1	RL18_SYNY3
225	7	2.2	122 1	RL18_SYNY3
226	7	2.2	122 1	RL18_SYNY3
227	7	2.2	122 1	RL18_SYNY3
228	7	2.2	122 1	RL18_SYNY3
229	7	2.2	122 1	RL18_SYNY3
230	7	2.2	122 1	RL18_SYNY3
231	7	2.2	122 1	RL18_SYNY3
232	7	2.2	122 1	RL18_SYNY3
233	7	2.2	122 1	RL18_SYNY3
234	7	2.2	122 1	RL18_SYNY3
235	7	2.2	122 1	RL18_SYNY3
236	7	2.2	122 1	RL18_SYNY3
237	7	2.2	122 1	RL18_SYNY3
238	7	2.2	122 1	RL18_SYNY3
239	7	2.2	122 1	RL18_SYNY3
240	7	2.2	122 1	RL18_SYNY3
241	7	2.2	122 1	RL18_SYNY3
242	7	2.2	122 1	RL18_SYNY3
243	7	2.2	122 1	RL18_SYNY3
244	7	2.2	122 1	RL18_SYNY3
245	7	2.2	122 1	RL18_SYNY3
246	7	2.2	122 1	RL18_SYNY3
247	7	2.2	122 1	RL18_SYNY3
248	7	2.2	122 1	RL18_SYNY3
249	7	2.2	122 1	RL18_SYNY3
250	7	2.2	122 1	RL18_SYNY3
251	7	2.2	122 1	RL18_SYNY3
252	7	2.2	122 1	RL18_SYNY3
253	7	2.2	122 1	RL18_SYNY3
254	7	2.2	122 1	RL18_SYNY3
255	7	2.2	122 1	RL18_SYNY3
256	7	2.2	122 1	RL18_SYNY3
257	7	2.2	122 1	RL18_SYNY3
258				

107	130	1	HIS3_PSESM	Q87u9 pseudomonas	180	6	2.2	201	1	LEUD_METEX	Q87p8 methylobact
108	134	6	HIS3_PSEAE	Q8uh7 pseudomonas	181	6	2.2	202	1	BIN3_STAU	P0384 staphylococ
109	134	6	Y652_MENIA	Q5868 methanococ	182	6	2.2	203	1	TRMB_MYCPU	Q88r4 mycoplasma
110	137	6	COF1_DICDI	P54706 dictyostell	183	6	2.2	204	1	RUVA_HABIN	P4632 haemophilus
111	138	6	RBFA_PSESM	Q87u96 pseudomonas	184	6	2.2	204	1	RUVA_VIBPA	Q87u8 vibrio para
112	143	6	PSAG_HORVU	Q00327 hordium vil	185	6	2.2	204	1	RUVA_VIBU	Q88r3 oceanobacti
113	143	6	R191_ARATH	Q95946 arabidopsis	186	6	2.2	206	1	THIE_OCBIR	Q87j2 yeastsia pe
114	143	6	R192_ARATH	Q9130 arabidopsis	187	6	2.2	207	1	KGUA_YERPE	Q87j2 yeastsia pe
115	144	6	USPF_ECOL6	Q87u3 escherichia	188	6	2.2	208	1	RLI3_SCHPO	P0155 schizosach
116	144	6	USPF_SALTU	Q8uh5 salmonella	189	6	2.2	208	1	RL4_MYCAA	P1015 mycoplasma
117	146	6	GLB1_ANABR	P02212 anadara bro	190	6	2.2	208	1	VPTI_MAIZE	P16976 zea mays (m
118	146	6	GLB1_SCATN	P02212 scapharca i	191	6	2.2	209	1	RNH2_CORGL	Q88r4 corynebact
119	146	6	GLB2_ANATR	P14394 anadara tra	192	6	2.2	210	1	ALKB_TRAPA	Q8358 t putative
120	147	6	RS5_CAMOE	Q9p1y8 campylabact	193	6	2.2	210	1	TRMB_MYCGE	P7559 mycoplasma
121	148	6	RL9_LISMO	Q8yar2 listeria mo	194	6	2.2	211	1	HET3_RADMG	Q8u6r1 radianthus
122	148	6	RL9_PSEBK	Q88d1 pseudomonas	195	6	2.2	211	1	MDCG_PSESM	Q87v9 pseudomonas
123	148	6	RL9_PSESM	Q87u46 pseudomonas	196	6	2.2	211	1	FLBE_CAUCR	Q04953 caulobacter
124	149	6	NAP5_ECOLI	P33333 escherichia	197	6	2.2	213	1	CYSE_STEAM	Q99w74 staphylococ
125	149	6	YID1_ECOLI	P31446 escherichia	198	6	2.2	215	1	PSB6_YEAST	P38524 saccharomyc
126	151	6	DTD_IACUA	Q9CJ92 lactococcus	199	6	2.2	215	1	YKJ7_YEAST	P28777 saccharomyc
127	151	6	RS9_MYCTU	Q06259 mycobacteri	200	6	2.2	218	1	TRKA_METUA	Q8505 methanococ
128	153	6	NDK3_SPIOL	P81768 spinacia ol	201	6	2.2	218	1	Y522_METUA	Q87942 methanococ
129	153	6	RS9_MYCLE	P40828 mycobacteri	202	6	2.2	220	1	Y064_ARCFU	Q30112 archaeoglob
130	155	6	CYNS_PSESM	Q885a6 pseudomonas	203	6	2.2	222	1	SFSA_THRAC	Q8hk4 thermoplas
131	155	6	REG1_PYRAB	Q9v2m0 pyrococcus	204	6	2.2	225	1	ATP6_LOCOMI	P44569 locusta mig
132	155	6	REG1_PYRHO	Q57818 pyrococcus	205	6	2.2	226	1	HAMI_CHUTE	Q8Kf6 chlorobium
133	156	6	BCCP_PSEAE	P37799 pseudomonas	206	6	2.2	226	1	INH_BPT4	P8808 bacterioph
134	157	6	GREB_SALTU	Q8z317 salmonella	207	6	2.2	228	1	V141_ARCFU	Q88r1 archaeoglob
135	157	6	GREB_SALTU	Q8z1j2 salmonella	208	6	2.2	229	1	NEP1_PYRHO	Q50081 pyrococcus
136	158	6	GREB_XANAC	Q8p1d9 xanthomonas	209	6	2.2	229	1	Y767_CABEL	Q11083 caenorhabd
137	158	6	GREB_XANCP	Q914d3 xanthomonas	210	6	2.2	231	1	Y7RE_BACSU	Q4392 bacillus su
138	158	6	GREB_ECOL6	Q8x723 escherichia	211	6	2.2	232	1	YVPH_CHICK	P01013 gallus gall
139	158	6	GREB_ECOLI	P30128 escherichia	212	6	2.2	232	1	YVPH_ECOLI	P76551 escherichia
140	160	6	TIMH_CABEL	Q2165 caenorhabd	213	6	2.2	234	1	NAGB_STIRP	Q99280 streptococ
141	161	6	COTF_BACSU	P2361 bacillus su	214	6	2.2	239	1	SFSA_AGRIS	P84295 agrobacteri
142	161	6	YB83_MENUA	Q58583 methanococ	215	6	2.2	241	1	CBID_SALTU	Q05550 salmonella
143	165	6	VEAE_BPP22	Q03547 bacterioph	216	6	2.2	242	1	TRPD_BACCA	Q05525 bacillus ca
144	165	6	YG25_HABIN	P44277 haemophilus	217	6	2.2	242	1	YU27_XANCP	Q8p6h3 xanthomonas
145	167	6	T2DE_YEAST	P11747 saccharomyc	218	6	2.2	242	1	YV51_XANCP	Q8p6h8 xanthomonas
146	173	6	YR15_CABEL	Q09355 caenorhabd	219	6	2.2	243	1	YVHO_SALTU	Q8x99 salmonella
147	175	6	OMLA_PSEFL	Q68564 pseudomonas	220	6	2.2	245	1	PHOS_FELCA	Q66374 colletoctric
148	176	6	RIMM_THEMA	Q9x1q4 thermocoga	221	6	2.2	245	1	PYRF_YERPE	P81666 felis silve
149	176	6	SSB_HABIN	P59930 haemophilus	222	6	2.2	245	1	YABS_BACSU	P75561 bacillus su
150	178	6	GRPE_RICPR	Q925t4 rickettsia	223	6	2.2	245	1	YVHO_ECO57	Q8x83 escherichia
151	178	6	YB16_SCHPO	Q10195 schizosach	224	6	2.2	245	1	YVHO_ECOLI	P77483 escherichia
152	180	6	IPYR_BUCAP	Q8k831 buchnera ap	225	6	2.2	246	1	RL7A_SCHPO	P17937 schizosach
153	180	6	YCBU_ECOLI	P75859 escherichia	226	6	2.2	249	1	LACR_STIAM	Q99873 staphylococ
154	182	6	IF3_SYMBL	Q8d1g8 synchococ	227	6	2.2	251	1	LACR_STIAM	P16644 staphylococ
155	182	6	PIPI_SYMBL	P71578 mycobacteri	228	6	2.2	251	1	SODE_MOUSE	Q09144 mus musculu
156	182	6	ORN_MYCTU	Q83c93 coxiella bu	229	6	2.2	252	1	NAGB_STIAM	Q99w40 staphylococ
157	183	6	SP2_YEAST	Q02521 saccharomyc	230	6	2.2	252	1	VG12_BPM2	Q84265 mycobacteri
158	185	6	T2_MOUSE	Q06666 mus musculu	231	6	2.2	254	1	HIS6_METUA	Q91v5 neisseria m
159	186	6	ENGB_MYCPU	Q984x1 mycoplasma	232	6	2.2	255	1	HIS6_NEIME	Q8K04 neisseria m
160	187	6	ISP2_VIBPA	Q87na5 vibrio para	233	6	2.2	255	1	PYRH_MYCLE	Q33045 mycobacteri
161	188	6	Y768_AQUAE	O66967 aquifex aeo	234	6	2.2	255	1	SSUB_ECOLI	P83053 escherichia
162	189	6	INAA_HUMAN	P05014 homo sapien	235	6	2.2	259	1	KSGA_MYCGE	P47701 m dimethyla
163	189	6	INAA_HUMAN	P0167 homo sapien	236	6	2.2	259	1	YK20_METUA	Q84426 methanococ
164	189	6	INAA_HUMAN	P32881 homo sapien	237	6	2.2	260	1	CLCA_PSEPU	P11451 pseudomonas
165	189	6	INAA_HUMAN	P0166 homo sapien	238	6	2.2	260	1	YM97_STRCO	Q91016 streptomyce
166	189	6	INAA_HUMAN	P05015 homo sapien	239	6	2.2	262	1	GLNH_BACST	P27676 bacillus st
167	189	6	INAA_HUMAN	P01571 homo sapien	240	6	2.2	265	1	PEN2_BACHD	Q8K808 caenorhabd
168	189	6	INAA_HUMAN	P01568 homo sapien	241	6	2.2	267	1	GNP1_CABEL	Q9xv2 caenorhabd
169	192	6	COMK_BACSU	P40396 bacillus su	242	6	2.2	268	1	TR1_DROME	P36188 drosophila
170	193	6	YD14_MYCTU	Q10622 mycobacteri	243	6	2.2	268	1	COB5_DEIRA	Q81v2 bacillus ce
171	194	6	Y667_DEIRA	Q97y6 detnoco	244	6	2.2	269	1	Y9C0_BACCA	Q81v2 bacillus an
172	196	6	YPI3_XANCP	Q8p1u5 xanthomonas	245	6	2.2	270	1	NLPA_ECO57	Q8Kx50 escherichia
173	198	6	YHMF_ECOLI	O06735 bacillus su	246	6	2.2	272	1	YHMF_ECOLI	P48464 escherichia
174	199	6	CYCO_BACSU	Q97f5 detnoco	247	6	2.2	272	1	DABI_ECOLI	P82210 rhizobium
175	199	6	Y670_DEIRA	O57213 haemophilus	248	6	2.2	273	1	Y12F_BPT4	P12225 bacterioph
176	200	6	RECR_HABIN	P44712 haemophilus	249	6	2.2	278	1	TRPA_RHTLO	Q88c6 rhizobium
177	200	6	RR4_GUTTH	P17072 guillardia	250	6	2.2	279	1		
178	200	6			251	6	2.2				
179	200	6			252	6	2.2				

253	6	2.2	280	1	OXA2_BACHD	Q9knb2 bacillus ha	326	6	2.2	335	1	RUVB_CAMTE	Q9pmt7 campylobact
254	6	2.2	284	1	EUTC_PROCLD	Q7nd23 phocorhabdu	327	6	2.2	336	1	DHOW_METTA	Q859m7 methanococc
255	6	2.2	285	1	ECHC_MYCTU	Q531e3 mycobacteri	328	6	2.2	336	1	GCP_BUCAL	P57166 buchnera ap
256	6	2.2	289	1	ATP6_PROMO	P21903 propionigen	329	6	2.2	336	1	HUTG_VIBVU	Q8da19 vibrio valn
257	6	2.2	291	1	YOHO_BACSU	P54513 bacillus su	330	6	2.2	336	1	HUTG_VIBVU	P60111 vibrio valn
258	6	2.2	292	1	DHVS_PYRAE	Q8z09 pyrobaculum	331	6	2.2	336	1	RUVB_HELPJ	Q9ams7 helicobacte
259	6	2.2	293	1	REFH_ECOLI	P27831 escherichia	332	6	2.2	337	1	RUCV_HELPJ	Q25699 helicobacte
260	6	2.2	294	1	NLPT_ECOLI	P39833 escherichia	333	6	2.2	337	1	RTCA_SUISO	Q97w04 sulfobus
261	6	2.2	294	1	PTA_THEMA	Q9x014 theomocoga	334	6	2.2	338	1	MDHM_HUYAN	P04926 homo sapien
262	6	2.2	296	1	NADC_RHOKU	P77938 rhodospirill	335	6	2.2	338	1	MDHM_MOUSE	P08242 mus musculu
263	6	2.2	296	1	NIFH_PLEBO	Q00240 plectonema	336	6	2.2	338	1	MDHM_PIG	P00346 sus scrofa
264	6	2.2	298	1	DAZL_MOUSE	Q64368 mus musculu	337	6	2.2	338	1	MDHM_RAT	P04636 rattus norv
265	6	2.2	298	1	MMSB_PSEAE	P28811 pseudomonas	338	6	2.2	339	1	NOAA_BACBD	Q9k9w9 bacillus ha
266	6	2.2	299	1	CSNS_SCHPO	Q94454 schizosacch	339	6	2.2	339	1	ADP_CHIKE	P31692 chlorella k
267	6	2.2	299	1	NANA_RHIME	Q92w00 rhizobium m	340	6	2.2	340	1	CEFA_BACUT	Q45729 bacillus th
268	6	2.2	299	1	SUCD_BACSU	P80865 bacillus su	341	6	2.2	340	1	COND_NEUCR	P14643 neurospora
269	6	2.2	300	1	ARGB_METUA	Q60382 methanococc	342	6	2.2	341	1	Y534_TREPA	O83545 treponema p
270	6	2.2	300	1	PSD_FUSNN	Q8r3f2 fusobacteri	343	6	2.2	342	1	ISIA_SYNP7	P15347 synechococc
271	6	2.2	302	1	RBGR_CHEVI	P25544 chromacium	344	6	2.2	342	1	ISIA_SYNP3	O55274 synechocyst
272	6	2.2	302	1	SUCD_STPAM	Q99um4 staphylococ	345	6	2.2	344	1	HRCA_STREN	O5442 streptococc
273	6	2.2	302	1	SUCD_STPAM	Q8n01 staphylococ	346	6	2.2	344	1	LFXD_SYNP3	Q95612 synechocyst
274	6	2.2	302	1	SUCD_STPAP	Q8cch4 staphylococ	347	6	2.2	345	1	UI94_HCVYA	P16800 human cytom
275	6	2.2	303	1	OSTP_PIG	P14287 sus scrofa	348	6	2.2	346	1	FMIR_GORCO	P79176 gorilla gor
276	6	2.2	303	1	Y367_RICPR	Q9zdg2 ticketstia	349	6	2.2	346	1	FMIR_MACMU	P79189 macaca mula
277	6	2.2	304	1	NODI_RHIS3	P72335 rhizobium s	350	6	2.2	346	1	FMIR_PANTR	P79241 pan troglod
278	6	2.2	305	1	E2BA_HUMAN	Q14332 homo sapien	351	6	2.2	346	1	FMIR_PONRY	P79235 pongo pygma
279	6	2.2	305	1	NODI_BRASS	Q9z313 bradyrhizob	352	6	2.2	347	1	HPT_MOUSE	O61646 mus musculu
280	6	2.2	306	1	NODI_BRAJA	P26050 bradyrhizob	353	6	2.2	347	1	HPT_MUSCR	O60574 mus caroli
281	6	2.2	307	1	YDCI_ECOLI	P77171 escherichia	354	6	2.2	347	1	HPT_MUSSA	Q62358 mus saxicol
282	6	2.2	308	1	MEIR_ECOLI	Q8x610 escherichia	355	6	2.2	347	1	HPT_PIG	Q88p57 sus scrofa
283	6	2.2	308	1	MEIR_ECOLI	P07623 escherichia	356	6	2.2	347	1	HPT_RAT	P06867 rattus norv
284	6	2.2	308	1	MEIR_ECOLI	Q8z1w1 salmonella	357	6	2.2	347	1	SELD_ECOLI6	O8fgy3 escherichia
285	6	2.2	308	1	MEIR_SALTI	P37413 salmonella	358	6	2.2	348	1	SELD_ECOLI	O14556 escherichia
286	6	2.2	308	1	RNH3_LISIN	Q92c10 listeria in	359	6	2.2	347	1	SELD_SALTI	O8x6f4 salmonella
287	6	2.2	308	1	RNH3_LISMO	Q9ydp5 listeria no	360	6	2.2	347	1	SELD_SALTY	O8xpy5 salmonella
288	6	2.2	309	1	ARCE_MYCPN	P78030 mycoplasma	361	6	2.2	347	1	SELD_SHIFL	P58393 shigella fl
289	6	2.2	309	1	CF35_MOUSE	Q9c226 mus musculu	362	6	2.2	348	1	HPTR_HUMAN	P00733 homo sapien
290	6	2.2	309	1	KHSE_SALTY	Q8xps5 salmonella	363	6	2.2	348	1	SELD_HAEIN	P49311 haemophilus
291	6	2.2	310	1	PPAC_STRMU	O68579 streptococc	364	6	2.2	348	1	XYLD_RHIOU	Q98d10 rhizobium l
292	6	2.2	310	1	TF2B_METH	O26971 methanobact	365	6	2.2	349	1	ASFP_VULVU	P53353 vulpes vulp
293	6	2.2	311	1	HPRK_MYCGS	P47331 mycoplasma	366	6	2.2	350	1	FMIR_HUMAN	P21462 homo sapien
294	6	2.2	312	1	CELF_PVYKA	O85330 pseudorabie	367	6	2.2	350	1	MPP_DEIRA	O97w55 deinozocuss
295	6	2.2	313	1	HEN3_SALTY	Q916q2 salmonella	368	6	2.2	352	1	FMIX_ECOLI	P10122 escherichia
296	6	2.2	313	1	MOA4_SULTO	Q977h1 sulfolobus	369	6	2.2	354	1	BCA4_ARATH	O91e06 arabidopsis
297	6	2.2	313	1	YDIR_BACSU	Q55025 bacillus su	370	6	2.2	355	1	NDI1_RHIME	O5618 rhizobium m
298	6	2.2	314	1	FATD_VIBAN	P37738 vibrio angu	371	6	2.2	356	1	GBA2_USITWA	P87033 ustulago ma
299	6	2.2	314	1	TYST_CANAL	P12461 candida alb	372	6	2.2	358	1	LAR2_RAT	Q924u0 rattus norv
300	6	2.2	316	1	CZCD_ALCOB	P13512 alcaligenes	373	6	2.2	359	1	DPO4_CLOPE	O8xk17 clostridium
301	6	2.2	316	1	CZCD_ALCSP	P94178 alcaligenes	374	6	2.2	359	1	PGS2_HUMAN	P07585 homo sapien
302	6	2.2	316	1	RAM2_YEAST	P29703 s protein f	375	6	2.2	359	1	YD56_AGR75	O8f6p1 agrobacteri
303	6	2.2	316	1	Y034_TREPA	O83077 treponema p	376	6	2.2	359	1	YOM3_CAEEL	Q09556 caenorhabdi
304	6	2.2	318	1	ATRA_ARATH	P92953 arabidopsis	377	6	2.2	360	1	LAR2_MOUSE	Q9j119 mus musculu
305	6	2.2	318	1	KHSE_VIBCB	Q9xk4 vibrio chol	378	6	2.2	360	1	PGS2_BOVIN	P21793 bos taurus
306	6	2.2	321	1	YB28_SYNP3	P72655 synechocyst	379	6	2.2	360	1	PGS2_CANPA	Q29393 canis famli
307	6	2.2	321	1	YDGT_SCHPO	Q10494 schizosacch	380	6	2.2	360	1	PGS2_PIG	Q9xsd8 sus scrofa
308	6	2.2	322	1	RADA_METVO	O73948 methanococc	381	6	2.2	360	1	PGS2_RABIT	Q88888 oryctolagus
309	6	2.2	322	1	TPIC_SPIOL	P48496 spirinacia ol	382	6	2.2	361	1	PGS2_SHEEP	Q9ctee ovis aries
310	6	2.2	322	1	Y711_METUA	Q58421 methanococc	383	6	2.2	362	1	AROB_BACHD	Q9xkb6 bacillus ha
311	6	2.2	323	1	OXS1_HUMAN	Q8n9j8 homo sapien	384	6	2.2	362	1	SERC_XYLFA	Q9pb19 xyella fas
312	6	2.2	325	1	CYSK_SOLTU	O81154 solanum tub	385	6	2.2	362	1	SERC_XYLFT	Q87bu0 xyella fas
313	6	2.2	326	1	TF2B_ARCFU	O28970 archaeoglob	386	6	2.2	366	1	PEL3_ASPEL	P41750 aspergillus
314	6	2.2	328	1	YIAO_HAEIN	P44992 haemophilus	387	6	2.2	368	1	PUR7_VIBCH	Q9kxte vibrio chol
315	6	2.2	329	1	HPT_CANPA	P19906 canis famli	388	6	2.2	369	1	DHDO_DICDI	P07670 dictyosteli
316	6	2.2	329	1	SUBI_SALTY	P02906 salmonella	389	6	2.2	369	1	GSEP_CAPAN	P80042 capsicum an
317	6	2.2	330	1	ODEB_BACSU	P37940 bacillus su	390	6	2.2	369	1	RADA_CERSY	O33748 cenarchaeum
318	6	2.2	332	1	AARP_WHEAT	Q02066 trititum ae	391	6	2.2	370	1	GIDA_BACST	P22816 bacillus st
319	6	2.2	332	1	MPGD_CLOAB	Q971d6 clostridium	392	6	2.2	374	1	TMU1_SULTO	Q97119 sulfolobus
320	6	2.2	332	1	MYOD_DROME	P22816 clostridia	393	6	2.2	377	1	KAPR_APICA	P13319 alveia cal
321	6	2.2	333	1	PLSX_BACSU	P71018 bacillus su	394	6	2.2	377	1	SEN2_YEAST	P16658 saccharomyc
322	6	2.2	334	1	ILYC_STAPB	O8c9g6 staphylococ	395	6	2.2	381	1	ME24_ASPEL	P46073 aspergillus
323	6	2.2	334	1	OPPF_ECOLI	P77737 escherichia	396	6	2.2	381	1	T10B_MOUSE	Q9qzmd mus musculu
324	6	2.2	334	1	OPPF_SALTY	P08007 salmonella	397	6	2.2	382	1	OVAN_CONYA	P19104 coturnix co
325	6	2.2	335	1	NDI2_RAIME	Q8ghn6 rhizobium m	398	6	2.2	384	1	OPGG_PSEBY	P20400 pseudomonas

399	6	2.2	384	1	PKR2_HUMAN	Q9NFJ6	homo sapien	472	6	2.2	451	1	ERG8_YEAST	P45521	saccharomyc
400	6	2.2	385	1	DIAC_HUMAN	Q01459	homo sapien	473	6	2.2	453	1	BRL3_HUMAN	Q9NF51	saccharomyc
401	6	2.2	385	1	OVAC_CHICK	P01012	gallus gall	474	6	2.2	457	1	MBL1_ECOLI	Q9NF55	homo sapien
402	6	2.2	387	1	ENGX_VIBCH	Q9XV18	vibrio chol	475	6	2.2	457	1	YC47_SYNY3	P37773	escherichia
403	6	2.2	387	1	NMT_DROME	061613	drosophila	476	6	2.2	461	1	NCB1_HUMAN	P45411	synecocyst
404	6	2.2	388	1	AMPX_VRENN	P45460	streptomyce	477	6	2.2	462	1	ALB3_ARATH	Q91B94	homo sapien
405	6	2.2	388	1	MSOX_STRESS	P40854	streptomyce	478	6	2.2	463	1	NAOX_METUA	Q80653	arabidopsis
406	6	2.2	388	1	OVAT_CHICK	P01014	gallus gall	479	6	2.2	463	1	SYG1_MIGR8	Q8D299	methanococc
407	6	2.2	390	1	DXR_FUSRN	Q8E522	fusobacteri	480	6	2.2	463	1	SYG1_STAM	Q99211	staphylococ
408	6	2.2	391	1	DXR_LEPIN	Q8E522	fusobacteri	481	6	2.2	463	1	SYG1_STAM	Q99211	staphylococ
409	6	2.2	391	1	XRCL_THETN	Q8E522	fusobacteri	482	6	2.2	464	1	SYG1_CLOAB	Q97062	staphylococ
410	6	2.2	392	1	DXR_CORGL	Q8E522	fusobacteri	483	6	2.2	464	1	6PGD_SALTY	Q74628	salmonella
411	6	2.2	392	1	GALI_CANFA	Q9CKX4	canis famli	484	6	2.2	468	1	EF1A_HYDAT	P15554	hydra atten
412	6	2.2	392	1	MP12_AMBAR	P28744	ambrosia ar	485	6	2.2	469	1	C39A_HUMAN	Q39Y15	homo sapien
413	6	2.2	394	1	FGK_BACHD	Q9X714	bacillus ha	486	6	2.2	470	1	SUFI_ECOLI	P26658	escherichia
414	6	2.2	395	1	CYSK_SCHPO	P87131	schizosacch	487	6	2.2	470	1	SUFI_ECOLI	P40739	salmonella
415	6	2.2	396	1	ATPB_ADIPA	Q03064	adiantum ra	488	6	2.2	472	1	PECK_TRYBB	P33735	trypanosoma
416	6	2.2	396	1	DEOB_CLOPE	P27759	clostridium	489	6	2.2	473	1	YACA_BACSU	P37260	bacillus su
417	6	2.2	396	1	MP11_AMBAR	P27759	clostridium	490	6	2.2	473	1	REBM_MOUSE	Q00473	salmonella
418	6	2.2	397	1	DXR_HAETN	P44055	haemophilus	491	6	2.2	475	1	NIEB_ANNAZ	Q38833	anabaena az
419	6	2.2	397	1	MP13_AMBAR	P27761	ambrosia ar	492	6	2.2	475	1	NIEB_ANNAZ	P20527	anabaena sp
420	6	2.2	398	1	MP12_AMBAR	P27760	ambrosia ar	493	6	2.2	477	1	YFUM_RHOCA	P01057	gallus gall
421	6	2.2	398	1	FGK_IACLA	Q9C1W1	lactococcus	494	6	2.2	477	1	ET52_CHICK	P26347	trypanosoma
422	6	2.2	400	1	AATA_RHIME	Q02635	thiobium m	495	6	2.2	479	1	VS12_TRYBB	P21842	lacaribe vi
423	6	2.2	402	1	PGK_CHLEP	P52178	brassica ol	496	6	2.2	482	1	VS12_TRYBB	P21842	lacaribe vi
424	6	2.2	402	1	ASSY_PSEAE	Q927M5	chlamydia p	497	6	2.2	483	1	SYP_MYCPN	P35382	mycoplasma
425	6	2.2	405	1	ASSY_PSEAE	Q927M5	chlamydia p	498	6	2.2	483	1	VS12_TRYBB	P21842	lacaribe vi
426	6	2.2	405	1	ASSY_PSEAE	P59604	pseudomonas	499	6	2.2	483	1	VS12_TRYBB	P21842	lacaribe vi
427	6	2.2	405	1	ASSY_PSEAE	P59604	pseudomonas	500	6	2.2	483	1	VS12_TRYBB	P21842	lacaribe vi
428	6	2.2	405	1	ASSY_PSEAE	P59604	pseudomonas	501	6	2.2	483	1	VS12_TRYBB	P21842	lacaribe vi
429	6	2.2	405	1	ASSY_PSEAE	P59604	pseudomonas	502	6	2.2	483	1	VS12_TRYBB	P21842	lacaribe vi
430	6	2.2	406	1	ASSY_PSEAE	P59604	pseudomonas	503	6	2.2	483	1	VS12_TRYBB	P21842	lacaribe vi
431	6	2.2	406	1	ASSY_PSEAE	P59604	pseudomonas	504	6	2.2	483	1	VS12_TRYBB	P21842	lacaribe vi
432	6	2.2	408	1	YF1N_ECOLI	Q09871	schizosacch	505	6	2.2	487	1	WRK1_ARATH	Q34137	arabidopsis
433	6	2.2	412	1	YAGT_SCHPO	Q09871	schizosacch	506	6	2.2	490	1	YHFE_ECOLI	P32121	escherichia
434	6	2.2	413	1	APGM_SUTLO	Q975B3	sulfolobus	507	6	2.2	490	1	YHFE_ECOLI	P32121	escherichia
435	6	2.2	414	1	CUTS_STRCO	Q03757	streptomyce	508	6	2.2	493	1	CPE1_YEAST	P31591	mesocricetu
436	6	2.2	414	1	K193_HUMAN	Q03757	streptomyce	509	6	2.2	494	1	NORB_BACSU	Q03552	bacillus su
437	6	2.2	414	1	K193_HUMAN	Q03757	streptomyce	510	6	2.2	495	1	IMDH_METUA	P78141	metanococc
438	6	2.2	414	1	K193_HUMAN	Q03757	streptomyce	511	6	2.2	496	1	IMDH_METUA	P78141	metanococc
439	6	2.2	415	1	NMTB_MYCBO	Q09443	xanthomonas	512	6	2.2	498	1	DHAM_LEITA	Q34137	leishmania
440	6	2.2	415	1	NMTB_MYCBO	Q09443	xanthomonas	513	6	2.2	498	1	DHAM_LEITA	Q34137	leishmania
441	6	2.2	417	1	ARPA_HUMAN	Q09443	xanthomonas	514	6	2.2	500	1	LYCB_TYCES	Q43578	lycopersico
442	6	2.2	417	1	GLAI_PSEAE	Q09443	xanthomonas	515	6	2.2	500	1	LYCB_TYCES	Q43578	lycopersico
443	6	2.2	418	1	GLAI_PSEAE	Q09443	xanthomonas	516	6	2.2	500	1	LYCB_TYCES	Q43578	lycopersico
444	6	2.2	418	1	NER3_MOUSE	Q09443	xanthomonas	517	6	2.2	500	1	LYCB_TYCES	Q43578	lycopersico
445	6	2.2	419	1	LMB2_YERPE	Q09443	xanthomonas	518	6	2.2	501	1	LYCB_TYCES	Q43578	lycopersico
446	6	2.2	420	1	ASSY_YEAST	Q09443	xanthomonas	519	6	2.2	501	1	LYCB_TYCES	Q43578	lycopersico
447	6	2.2	420	1	ASSY_YEAST	Q09443	xanthomonas	520	6	2.2	501	1	LYCB_TYCES	Q43578	lycopersico
448	6	2.2	420	1	GLYA_PASMU	P52768	saccharomyc	521	6	2.2	501	1	LYCB_TYCES	Q43578	lycopersico
449	6	2.2	420	1	GLYA_PASMU	P52768	saccharomyc	522	6	2.2	501	1	LYCB_TYCES	Q43578	lycopersico
450	6	2.2	421	1	GLYA_PASMU	P52768	saccharomyc	523	6	2.2	501	1	LYCB_TYCES	Q43578	lycopersico
451	6	2.2	421	1	GLYA_PASMU	P52768	saccharomyc	524	6	2.2	501	1	LYCB_TYCES	Q43578	lycopersico
452	6	2.2	421	1	GLYA_PASMU	P52768	saccharomyc	525	6	2.2	501	1	LYCB_TYCES	Q43578	lycopersico
453	6	2.2	421	1	GLYA_PASMU	P52768	saccharomyc	526	6	2.2	501	1	LYCB_TYCES	Q43578	lycopersico
454	6	2.2	421	1	GLYA_PASMU	P52768	saccharomyc	527	6	2.2	501	1	LYCB_TYCES	Q43578	lycopersico
455	6	2.2	421	1	GLYA_PASMU	P52768	saccharomyc	528	6	2.2	501	1	LYCB_TYCES	Q43578	lycopersico
456	6	2.2	421	1	GLYA_PASMU	P52768	saccharomyc	529	6	2.2	501	1	LYCB_TYCES	Q43578	lycopersico
457	6	2.2	421	1	GLYA_PASMU	P52768	saccharomyc	530	6	2.2	501	1	LYCB_TYCES	Q43578	lycopersico
458	6	2.2	421	1	GLYA_PASMU	P52768	saccharomyc	531	6	2.2	501	1	LYCB_TYCES	Q43578	lycopersico
459	6	2.2	421	1	GLYA_PASMU	P52768	saccharomyc	532	6	2.2	501	1	LYCB_TYCES	Q43578	lycopersico
460	6	2.2	421	1	GLYA_PASMU	P52768	saccharomyc	533	6	2.2	501	1	LYCB_TYCES	Q43578	lycopersico
461	6	2.2	421	1	GLYA_PASMU	P52768	saccharomyc	534	6	2.2	501	1	LYCB_TYCES	Q43578	lycopersico
462	6	2.2	421	1	GLYA_PASMU	P52768	saccharomyc	535	6	2.2	501	1	LYCB_TYCES	Q43578	lycopersico
463	6	2.2	421	1	GLYA_PASMU	P52768	saccharomyc	536	6	2.2	501	1	LYCB_TYCES	Q43578	lycopersico
464	6	2.2	421	1	GLYA_PASMU	P52768	saccharomyc	537	6	2.2	501	1	LYCB_TYCES	Q43578	lycopersico
465	6	2.2	421	1	GLYA_PASMU	P52768	saccharomyc	538	6	2.2	501	1	LYCB_TYCES	Q43578	lycopersico
466	6	2.2	421	1	GLYA_PASMU	P52768	saccharomyc	539	6	2.2	501	1	LYCB_TYCES	Q43578	lycopersico
467	6	2.2	421	1	GLYA_PASMU	P52768	saccharomyc	540	6	2.2	501	1	LYCB_TYCES	Q43578	lycopersico
468	6	2.2	421	1	GLYA_PASMU	P52768	saccharomyc	541	6	2.2	501	1	LYCB_TYCES	Q43578	lycopersico
469	6	2.2	421	1	GLYA_PASMU	P52768	saccharomyc	542	6	2.2	501	1	LYCB_TYCES	Q43578	lycopersico
470	6	2.2	421	1	GLYA_PASMU	P52768	saccharomyc	543	6	2.2	501	1	LYCB_TYCES	Q43578	lycopersico
471	6	2.2	421	1	GLYA_PASMU	P52768	saccharomyc	544	6	2.2	501	1	LYCB_TYCES	Q43578	lycopersico

545	6	2.2	525	1	OPGG_PSEAE	Q9huas	pseudomonas	618	6	2.2	672	1	Y030_VIBVY	Q7utis	vibrio vuln
546	6	2.2	529	1	YEJF_ECOLI	P33916	escherichia	619	6	2.2	673	1	TF2B_METJA	O58192	methanococ
547	6	2.2	531	1	PYRG_SULTO	O97669	sulfobolus	620	6	2.2	682	1	RECG_BACSU	O34942	bacillus su
548	6	2.2	533	1	MUTL_THBAQ	P96082	thermus aqu	621	6	2.2	688	1	ERG_CLOPE	O8xhs1	clostridium
549	6	2.2	542	1	THS_METJA	O58405	methanococ	622	6	2.2	689	1	KR1B_RAT	O88658	rattus norv
550	6	2.2	543	1	CH60_MYCPN	P78012	mycoplasma	623	6	2.2	694	1	HAAL_YEAST	O12753	saccharomyc
551	6	2.2	543	1	TMMD_MYCPN	P87940	buchnera ap	624	6	2.2	697	1	ERG_RICBE	O8Rtbo	rikkettsia
552	6	2.2	544	1	FLKG_BUDAP	O8x9k0	glycine max	625	6	2.2	697	1	ERG_RICBE	O8Rtbo	rikkettsia
553	6	2.2	544	1	GSHC_SOYBN	P48640	arabidopsis	626	6	2.2	699	1	ERG_RICCN	O9t193	rikkettsia
554	6	2.2	544	1	Y290_ARATH	O04212	arabidopsis	627	6	2.2	699	1	ERG_RICCN	O8Rtba	rikkettsia
555	6	2.2	545	1	MUTL_THETH	O9454	thermus the	628	6	2.2	699	1	ERG_RICPE	O8Rtba	rikkettsia
556	6	2.2	547	1	CH60_LBGRN	P26878	legionella	629	6	2.2	699	1	ERG_RICPE	O8Rtba	rikkettsia
557	6	2.2	547	1	IF37_MOUSE	O70194	mus musculu	630	6	2.2	699	1	ERG_RICPE	O8Rtba	rikkettsia
558	6	2.2	548	1	IF37_MOUSE	O15371	homo sapien	631	6	2.2	699	1	ERG_RICPE	O8Rtba	rikkettsia
559	6	2.2	548	1	IF37_MOUSE	O15371	homo sapien	632	6	2.2	699	1	ERG_RICPE	O8Rtba	rikkettsia
560	6	2.2	549	1	Y4KD_RHISN	P55524	rhizobium s	633	6	2.2	699	1	ERG_RICPE	O8Rtba	rikkettsia
561	6	2.2	553	1	YIDE_SALTI	O82217	salmonella	634	6	2.2	699	1	ERG_RICPE	O8Rtba	rikkettsia
562	6	2.2	553	1	YIDE_SALTI	O82217	salmonella	635	6	2.2	699	1	ERG_RICPE	O8Rtba	rikkettsia
563	6	2.2	554	1	Y514_SYNY3	O54370	synchocyst	636	6	2.2	702	1	DXK4_MOUSE	O8Rtba	rikkettsia
564	6	2.2	555	1	PGMU_ASPOR	P57749	aspergillus	637	6	2.2	702	1	DXK4_MOUSE	O8Rtba	rikkettsia
565	6	2.2	556	1	HCP_THIPE	P96095	thiobacilli	638	6	2.2	704	1	PEB1_ORYSA	O8Rtba	rikkettsia
566	6	2.2	556	1	SYO_VIBCH	O9Kba6	vibrio chol	639	6	2.2	710	1	COG2_DROME	O8Rtba	rikkettsia
567	6	2.2	559	1	HUTU_BACHD	O9Kba5	babillus ha	640	6	2.2	710	1	NECE_HYDAT	O8Rtba	rikkettsia
568	6	2.2	564	1	DPY4_RAT	O62551	rattus norv	641	6	2.2	710	1	NECE_HYDAT	O8Rtba	rikkettsia
569	6	2.2	572	1	DPY4_BOVIN	O02675	bos taurus	642	6	2.2	713	1	DXK4_RAT	O8Rtba	rikkettsia
570	6	2.2	572	1	DPY4_CHICK	O90635	gallus gall	643	6	2.2	714	1	YEB7_YEAST	O8Rtba	rikkettsia
571	6	2.2	572	1	DPY4_CHICK	O16555	homo sapien	644	6	2.2	717	1	CY4B_HUMAN	O8Rtba	rikkettsia
572	6	2.2	572	1	DPY4_CHICK	O08553	mus musculu	645	6	2.2	721	1	PRTB_HSVJ7	O8Rtba	rikkettsia
573	6	2.2	572	1	DPY4_CHICK	O08553	mus musculu	646	6	2.2	724	1	DXK4_HUMAN	O8Rtba	rikkettsia
574	6	2.2	572	1	DPY4_CHICK	O14531	homo sapien	647	6	2.2	724	1	DXK4_HUMAN	O8Rtba	rikkettsia
575	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	648	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
576	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	649	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
577	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	650	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
578	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	651	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
579	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	652	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
580	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	653	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
581	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	654	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
582	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	655	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
583	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	656	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
584	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	657	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
585	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	658	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
586	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	659	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
587	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	660	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
588	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	661	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
589	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	662	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
590	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	663	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
591	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	664	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
592	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	665	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
593	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	666	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
594	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	667	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
595	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	668	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
596	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	669	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
597	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	670	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
598	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	671	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
599	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	672	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
600	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	673	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
601	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	674	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
602	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	675	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
603	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	676	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
604	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	677	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
605	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	678	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
606	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	679	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
607	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	680	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
608	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	681	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
609	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	682	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
610	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	683	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
611	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	684	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
612	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	685	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
613	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	686	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
614	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	687	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
615	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	688	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
616	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	689	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia
617	6	2.2	574	1	VGLE_HRSVA	O35098	mus musculu	690	6	2.2	725	1	TAP1_MOUSE	O8Rtba	rikkettsia

691	2.2	813	1	TIRE_ECOLI	Q47281	eschericchia	764	1111	1	NALS_MOUSE	Q911m5	mus musculus
692	2.2	815	1	GYRE_MYXXA	Q33167	myxococcus	765	1126	1	HEM_DROME	P55162	drosophila
693	2.2	819	1	LON_CHLBN	Q929f4	chlamydia p	766	1169	1	SUV3_YEAST	P20183	saccharomyc
694	2.2	824	1	DROT_METVO	P52025	methanococc	767	1176	1	VPS8_YEAST	P39702	saccharomyc
695	2.2	826	1	SSP2_PLAYO	O01443	plasmodium	768	1194	1	DPOL_VZVD	P09282	varicella-z
696	2.2	827	1	CADR_RAT	P54581	rattus norv	769	1194	1	KANK_HUMAN	O14678	homo sapien
697	2.2	829	1	IF2_HABIN	P43423	haemophilus	770	1210	1	ICEN_PSEFI	P09815	pseudomonas
698	2.2	830	1	EF2_DICDI	P15112	dictyostell	771	1238	1	DPOA_ORISA	P54099	mus musculus
699	2.2	830	1	MCW2_SCHPO	O13465	cryptococcu	772	1243	1	POLS_SINDO	P27285	coryza sativ
700	2.2	832	1	UREA_CRYNE	017392	caenorhabdi	773	1245	1	POLS_SINDO	P27285	coryza sativ
701	2.2	833	1	UREA_CAEEL	O13465	cryptococcu	774	1245	1	POLS_SINDO	P27285	coryza sativ
702	2.2	835	1	UREA_SCHPO	O17392	caenorhabdi	775	1245	1	POLS_SINDO	P27285	coryza sativ
703	2.2	836	1	CEG1_HUMAN	Q09910	homo sapien	776	1257	1	Y211_HUMAN	TOXA_PASMU	
704	2.2	836	1	RPOC_CHLVU	P56300	chlorella v	777	1267	1	MIP1_SCHPO	EF2_METXA	
705	2.2	837	1	UBPA_DICDI	P54201	dictyostell	778	1313	1	MRP4_HUMAN	EF2_METXA	
706	2.2	840	1	NLGI_HUMAN	Q8n2q7	homo sapien	779	1325	1	MRP4_HUMAN	EF2_METXA	
707	2.2	843	1	NLGI_MOUSE	Q9n910	mus musculus	780	1360	1	CING_XENLA	NEOI_RAT	
708	2.2	843	1	NLGI_MOUSE	Q62765	rattus norv	781	1377	1	NEOI_RAT	ICP4_HSYWG	
709	2.2	849	1	AKA4_MOUSE	Q60662	mus musculus	782	1415	1	ICP4_HSYWG	SMC4_YEAST	
710	2.2	850	1	OPGH_EKMC	Q9f495	erwinia citr	783	1418	1	NEOI_CHICK	ADP1_MYCGE	
711	2.2	853	1	ADG_USTMA	Q99128	ustillago ma	784	1443	1	ADP1_MYCGE	ADP1_MYCGE	
712	2.2	859	1	YD48_MYCTU	Q11018	mycobacteri	785	1444	1	ADP1_MYCGE	ADP1_MYCGE	
713	2.2	861	1	OPGH_PSEAE	Q9hu46	pseudomonas	786	1451	1	EM30_ARAYH	NEOI_HUMAN	
714	2.2	868	1	MCW2_YEAST	P29469	saccharomyc	787	1451	1	NEOI_HUMAN	BCL9_DROME	
715	2.2	869	1	SVL_SYNV3	P73374	synchocyst	788	1461	1	NEOI_HUMAN	BCL9_DROME	
716	2.2	876	1	KCH8_MOUSE	P59111	mus musculus	789	1469	1	NEOI_HUMAN	BCL9_DROME	
717	2.2	876	1	PRP2_YEAST	P20095	saccharomyc	790	1486	1	NEOI_HUMAN	BCL9_DROME	
718	2.2	879	1	MANB_BOVIN	Q29444	bos taurus	791	1493	1	NEOI_HUMAN	BCL9_DROME	
719	2.2	879	1	MANB_CAPII	Q95327	capra hircu	792	1510	1	NEOI_HUMAN	BCL9_DROME	
720	2.2	879	1	PUP3_YEAST	Q07807	saccharomyc	793	1516	1	NEOI_HUMAN	BCL9_DROME	
721	2.2	880	1	SYA_CAUCR	Q9A5C1	caulobacter	794	1636	1	SN24_HUMAN	FN37_YEAST	
722	2.2	880	1	YE21_ARCFU	Q28851	archaeoglob	795	1647	1	NEOI_HUMAN	FN37_YEAST	
723	2.2	882	1	Y890_MYCTU	O10550	mycobacteri	796	1679	1	NEOI_HUMAN	FN37_YEAST	
724	2.2	882	1	Y914_MYCBO	P59969	mycobacteri	797	1690	1	KE1A_HUMAN	FUR2_DROME	
725	2.2	882	1	YBAH_SCHPO	O42308	schizosacch	798	1695	1	KE1A_HUMAN	FUR2_DROME	
726	2.2	885	1	IF2_SHEON	Q8eh15	sheareella	799	1699	1	KE1A_HUMAN	FUR2_DROME	
727	2.2	887	1	MTF_BOVIN	P55156	bos taurus	799	1699	1	KE1A_HUMAN	FUR2_DROME	
728	2.2	887	1	MTF_BOVIN	P55156	bos taurus	800	1807	1	TSC2_HUMAN	TSC2_RAT	
729	2.2	888	1	SYA_METXA	O57984	methanococc	801	1809	1	TSC2_RAT	TSC2_RAT	
730	2.2	889	1	SECA_BORBU	P18751	xenopus lae	802	1816	1	KE1B_HUMAN	KE1B_MOUSE	
731	2.2	896	1	SECA_RICEP	O27497	borrelia bu	803	1816	1	KE1B_HUMAN	KE1B_MOUSE	
732	2.2	906	1	ITTH_MOUSE	Q92CA7	ricicetia	804	1887	1	FA52_YEAST	FA52_YEAST	
733	2.2	907	1	NUOG_ECO57	Q61702	mus musculus	805	1938	1	MSYD_SCHPO	BGS4_SCHPO	
734	2.2	907	1	NUOG_ECO57	Q8cx22	eschericchia	806	1955	1	OTOF_HUMAN	OTOF_HUMAN	
735	2.2	907	1	NUOG_ECOL6	Q8efj9	eschericchia	807	1955	1	OTOF_HUMAN	OTOF_HUMAN	
736	2.2	907	1	NUOG_ECOL1	O8fif9	eschericchia	808	1955	1	OTOF_HUMAN	OTOF_HUMAN	
737	2.2	907	1	NUOG_SALTY	P33900	salmonella	809	2116	1	RRPL_SYNV	GLT1_YEAST	
738	2.2	907	1	NUOG_SALTY	O7uc56	shigella fl	810	2144	1	GLT1_YEAST	GLT1_YEAST	
739	2.2	911	1	ITTH_HUMAN	P19827	homo sapien	811	2194	1	POLG_HERO1	POLG_HERO1	
740	2.2	914	1	ITTH_MESAU	P97278	mesocricetu	812	2223	1	CCAT_HUMAN	CCAT_HUMAN	
741	2.2	920	1	UBI1_HUMAN	P51784	homo sapien	813	2282	1	POLN_SFV	POLN_SFV	
742	2.2	928	1	MAV4_SCHCO	P37935	schizophyll	814	2284	1	POLN_SFV	POLN_SFV	
743	2.2	941	1	A2A2_RAT	O18484	rattus norv	815	2431	1	ABEC2_HUMAN	ABEC2_HUMAN	
744	2.2	941	1	DNAB_RHOMR	P30477	rhodothermu	816	2431	1	ABEC2_HUMAN	ABEC2_HUMAN	
745	2.2	946	1	IP3L_HUMAN	P27987	homo sapien	817	2468	1	ABEC2_HUMAN	ABEC2_HUMAN	
746	2.2	950	1	Y136_HUMAN	O14149	homo sapien	818	2468	1	ABEC2_HUMAN	ABEC2_HUMAN	
747	2.2	952	1	IF41_YEAST	P39935	saccharomyc	819	2594	1	NAPE_HUMAN	NAPE_HUMAN	
748	2.2	960	1	DIG1_DROME	P31007	drosophila	820	2696	1	CVAA_USTMA	CVAA_USTMA	
749	2.2	971	1	RECK_HUMAN	O95980	homo sapien	821	2842	1	THES_DROVI	THES_DROVI	
750	2.2	977	1	VAG9_SCHPO	O09872	schizosacch	822	2843	1	APC_HUMAN	APC_HUMAN	
751	2.2	1005	1	DPOL_VARY	P33793	variola vir	823	2843	1	APC_HUMAN	APC_HUMAN	
752	2.2	1006	1	DPOL_VACC	P20508	vaccinia vi	824	2869	1	RAP1_PLAYB	RAP1_PLAYB	
753	2.2	1006	1	DPOL_VACC	P20508	vaccinia vi	825	2869	1	RAP1_PLAYB	RAP1_PLAYB	
754	2.2	1012	1	IF2C_PHAUV	P06856	vaccinia vi	826	3172	1	ERYV_SACER	ERYV_SACER	
755	2.2	1021	1	AI41_HORSE	P57997	phaseolus v	827	3210	1	ERYV_SACER	ERYV_SACER	
756	2.2	1029	1	END1_YEAST	P18907	equus caball	828	3491	1	ACVS_NOCIA	ACVS_NOCIA	
757	2.2	1045	1	CUSA_ECO57	P12868	saccharomyc	829	3649	1	ACVS_NOCIA	ACVS_NOCIA	
758	2.2	1047	1	CUSA_ECOL6	O8f436	eschericchia	830	3703	1	ABF1_HUMAN	ABF1_HUMAN	
759	2.2	1053	1	SLPM_BACBR	P38054	eschericchia	831	3898	1	POLG_HCVB	POLG_HCVB	
760	2.2	1082	1	RPOB_EUGER	P06546	escherichia	832	4092	1	DYHC_YEAST	DYHC_YEAST	
761	2.2	1093	1	AP17_HUMAN	P23579	euglena gra	833	4289	1	TENX_HUMAN	TENX_HUMAN	
762	2.2	1102	1	KCH8_RAT	P55198	homo sapien	834	4349	1	TAT2_HUMAN	TAT2_HUMAN	
763	2.2	1102	1	RPOB_AGABT	O9uwa8	rattus norv	835	4351	1	FAT2_RAT	FAT2_RAT	
					P33539	agaricus bi	836					

837	6	2.2	4473	1	PLEI CRIGR	09155 cricetus	910	1.8	68	1	YB10 HALN1	P1704 halobacteri
838	6	2.2	4590	1	FATH HUMAN	014517 homo sapien	911	1.8	69	1	ACP RHOSH	P1784 rhodobacter
839	6	2.2	4639	1	DYHC DROME	P37276 drosophila	912	1.8	69	1	GBGT BOVIN	P5154 bos taurus
840	6	2.2	4664	1	PLEI_HUMAN	015149 homo sapien	913	1.8	69	1	YCH_RHISN	P53390 rhizobium s
841	6	2.2	4687	1	PLEI_RAT	P30427 rattus norv	914	1.8	70	1	IBY7 MYCJM	Q9XCD6 mycobacteri
842	6	2.2	5065	1	EPPL_HUMAN	P58107 homo sapien	915	1.8	70	1	INS_TORMA	P12705 torpeda mar
843	6	2.2	5120	1	PCLO CHICK	Q9AP36 gallus gall	916	1.8	70	1	R37A SULSO	Q97393 sulfobius
844	6	2.2	5127	1	RY44 DROME	024496 drosophila	917	1.8	70	1	RLJ1 BIULO	Q89396 bifidobacte
845	6	2.2	5147	1	FAT DROME	P33450 drosophila	918	1.8	70	1	YF73 META	Q58974 methanococc
846	6	2.2	5171	1	BPEA HUMAN	094833 homo sapien	919	1.8	71	1	GVPA_PTAG	Q97300 planktontri
847	6	2.2	6486	1	TYCC BREPA	030405 b tyrocidin	920	1.8	71	1	VG6_SPV4	P11338 spiroplasma
848	6	2.2	6548	1	EPPL_MOUSE	Q80W0 mus muscula	921	1.8	71	1	VTB2 XENIA	P19011 xenopus lae
849	6	2.2	7073	1	RIAB CVASA	P56411 h replicase	922	1.8	72	1	IF1 BIFLO	Q89327 bifidobacte
850	6	2.2	7389	1	BPAL MOUSE	Q91206 mus musculi	923	1.8	73	1	MPT5 AMBR	P10414 ambrosia tr
851	6	1.8	23	1	CR41 LITCE	P56243 litorea cae	924	1.8	74	1	ATP9 LYCES	P60117 lycopersico
852	5	1.8	23	1	CR43 LITCE	P56244 litorea cae	925	1.8	74	1	ATP9 MARZE	P00840 zea mays (m
853	5	1.8	23	1	BRIC_RANSP	P82906 rana spheno	926	1.8	74	1	ATP9_OENBI	P26855 marchantia
854	5	1.8	24	1	FEDG_AMTME	P80707 amycolatops	927	1.8	74	1	ATP9_PETRY	P60115 petunia hyb
855	5	1.8	25	1	ALR_PSEPL	P33162 pseudomonas	928	1.8	74	1	ATP9_SOLTY	P60114 solanum tub
856	5	1.8	26	1	FORA_MERTM	P80900 methanobact	930	1.8	74	1	ATP9_TOBAC	P60116 nicotiana t
857	5	1.8	26	1	YFHA_KLEOX	P21710 klebsiella	931	1.8	74	1	ATP9_WHEAT	P13547 triticum ae
858	5	1.8	28	1	YAF9_ARCFU	029184 archaeoglob	932	1.8	74	1	RK14_OENAM	P42340 oenothera s
859	5	1.8	30	1	CYOT_VIOOD	P58439 viola odora	933	1.8	75	1	ATP1_RHOFU	P19704 rhodospirill
861	5	1.8	33	1	CECC_HELVI	P83413 heliothis v	934	1.8	75	1	HOLI_BPRIT	Q38134 bacterioph
862	5	1.8	35	1	Y812_PASMT	Q9CMK8 pasteurella	935	1.8	75	1	ACP ANAVA	P24022 lactobacill
863	5	1.8	42	1	RL11_STRGB	P52661 streptomyces	936	1.8	76	1	ACP_CLOPE	P20803 anabaena va
864	5	1.8	43	1	ACP PHOHR	Q97423 photobacter	937	1.8	76	1	ACP_HAEIN	Q87197 clostridium
865	5	1.8	45	1	CSPR_KLEPN	Q48493 klebsiella	938	1.8	76	1	ACP_LEUMU	P43709 haemophilus
866	5	1.8	46	1	YBBO_BACSU	034568 bacillus su	939	1.8	76	1	ACP_OCELI	P80920 leucocithrix
867	5	1.8	48	1	YK82_ARCFU	Q28197 archaeoglob	940	1.8	76	1	ACP_PASMT	Q80922 oceanospiri
868	5	1.8	49	1	R333_LACLA	Q9CAV5 lactococcus	941	1.8	76	1	ACP_THENT	Q87452 agrobacteri
869	5	1.8	49	1	Y5K8_SSV1	P20204 sulfobius	942	1.8	76	1	ACP_VIBPA	Q97145 pasteurella
870	5	1.8	51	1	ME23_EUPRA	P58547 euploies ra	943	1.8	76	1	TEGP_HSVPA	Q87337 vibrio para
871	5	1.8	51	1	RL33_ABRPE	Q97472 aeropyrum p	944	1.8	76	1	ACP_PSESM	P18554 equine hearp
872	5	1.8	51	1	RL33_SUITO	Q97141 sulfobius	945	1.8	77	1	ACP_RHILE	P57798 helicobacte
873	5	1.8	54	1	PERA_SULME	P81433 sulfobius	946	1.8	77	1	ACP_AGRIS	Q54439 pseudomonas
874	5	1.8	55	1	ATP8_POLOR	Q95912 polypterus	947	1.8	77	1	ACP_CIOAB	Q87452 agrobacteri
875	5	1.8	55	1	COPE_STRAU	P25921 staphylococ	948	1.8	77	1	ACP_COMTE	Q97145 pasteurella
876	5	1.8	57	1	MAJE_PSESI	P80693 pseudomonas	949	1.8	77	1	ACP_ECOLI	P55337 vibrio para
877	5	1.8	57	1	RL33_ANASP	Q87475 archaeoglob	950	1.8	77	1	ACP_HAEDU	P18554 equine hearp
878	5	1.8	59	1	YA90_ARCFU	Q29170 archaeoglob	951	1.8	77	1	ACP_PSESM	P57798 helicobacte
879	5	1.8	60	1	SERL_DROME	Q9VEW4 drosophila	952	1.8	77	1	ACP_RHILE	Q54439 pseudomonas
881	5	1.8	60	1	YAB7_ARCFU	029178 archaeoglob	953	1.8	77	1	ACP_RHIME	Q87452 agrobacteri
882	5	1.8	61	1	AKH2_LOCOM	P08379 locusta mig	954	1.8	77	1	ACP_SHEON	Q97145 pasteurella
883	5	1.8	61	1	ASSY_LACLC	Q59491 lactococcus	955	1.8	77	1	ACP_SYNY3	P20803 anabaena va
884	5	1.8	61	1	YAL2_ARCFU	029173 archaeoglob	956	1.8	77	1	ACP_VIBCH	Q87337 vibrio para
885	5	1.8	61	1	YAL2_ARCFU	028071 archaeoglob	957	1.8	77	1	ACP_VIBMA	Q87337 vibrio para
886	5	1.8	62	1	YA95_ARCFU	Q29170 archaeoglob	958	1.8	77	1	ACP_VIBMA	Q87337 vibrio para
887	5	1.8	62	1	CE3E_HYPCU	P50722 hypantria	959	1.8	77	1	ACP_YERPE	Q87337 vibrio para
888	5	1.8	63	1	CE3E_HYPCU	P50723 hypantria	960	1.8	77	1	IF1C_ANTMA	Q87337 vibrio para
889	5	1.8	63	1	DI07_HUMAN	Q81877 homo sapien	961	1.8	77	1	IF1C_BOROF	Q87337 vibrio para
890	5	1.8	63	1	RL29_NEIMA	Q91X24 neisseria m	962	1.8	77	1	IF1C_CORMA	Q87337 vibrio para
891	5	1.8	63	1	RPO9_MYXVA	Q9GDP6 myxoma viru	963	1.8	77	1	IF1C_GAREL	Q87337 vibrio para
892	5	1.8	63	1	RPO9_SFVKA	Q9G821 Shope fibro	964	1.8	77	1	IF1C_LEUPR	Q87337 vibrio para
893	5	1.8	63	1	YA84_ARCFU	029181 archaeoglob	965	1.8	77	1	IF1C_SANCA	Q87337 vibrio para
894	5	1.8	64	1	RL35_CHLCV	Q82243 chlamydomon	966	1.8	77	1	TAC1_TACTR	Q87337 vibrio para
895	5	1.8	64	1	RL35_CHLTR	084C41 chlamydia t	967	1.8	77	1	TAC2_TACTR	Q87337 vibrio para
896	5	1.8	64	1	Y314_ARCFU	029931 archaeoglob	968	1.8	77	1	Y188_ARCFU	Q87337 vibrio para
897	5	1.8	64	1	Y316_ARCFU	029929 archaeoglob	969	1.8	77	1	Y317_ARCFU	Q87337 vibrio para
898	5	1.8	64	1	Y474_ARCFU	029189 archaeoglob	970	1.8	77	1	ACP_AZOBR	Q87337 vibrio para
899	5	1.8	65	1	RL35_ARCFU	P57327 buchnera ap	971	1.8	78	1	ACP_BRMVE	Q87337 vibrio para
900	5	1.8	65	1	RL35_BUCBP	P49240 buchnera ap	972	1.8	78	1	ACP_NEIVA	Q87337 vibrio para
901	5	1.8	65	1	RS27_THEAC	Q89E99 buchnera ap	973	1.8	78	1	ACP_RHILLO	Q87337 vibrio para
902	5	1.8	65	1	RS27_THEAC	Q91X22 thermoplasma	974	1.8	78	1	ACP_XANAC	Q87337 vibrio para
903	5	1.8	65	1	SECE_THEMEA	P397P1 thermoplasma	975	1.8	78	1	ACP_XANAL	Q87337 vibrio para
904	5	1.8	65	1	Y319_ARCFU	029926 archaeoglob	976	1.8	78	1	COX8_YEAST	Q87337 vibrio para
905	5	1.8	66	1	R24E_HALMA	P14116 halocaula	977	1.8	78	1	FRIL_PIG	Q87337 vibrio para
906	5	1.8	66	1	YXCD_BACSU	Q8XGGO bacillus su	978	1.8	78	1	MULI_ECOLI	P02937 escherichia
907	5	1.8	67	1	RPO2_RALSO	Q8XGGO raietonia s	979	1.8	78	1	MULI_ERWAM	P02939 erwania amy
908	5	1.8	67	1			980	1.8	78	1		
909	5	1.8	67	1			981	1.8	78	1		
910	5	1.8	68	1			982	1.8	78	1		

983	5	1.8	78	1	MULTI-MORNO	P02940 morganella
984	5	1.8	78	1	MULTI-PROMI	P09461 proteus mir
985	5	1.8	79	1	ACPI-RALSO	Q8Y0J1 talstonia s
986	5	1.8	79	1	ACPI-BUCAP	Q8X9J4 buchiera ap
987	5	1.8	79	1	ACPI-XYLFA	Q9P6I5 xyella fas
988	5	1.8	79	1	EXTS-LACIA	Q9CBH3 lactococcus
989	5	1.8	79	1	YIF2-BACAA	Q81Y08 bacillus an
990	5	1.8	79	1	YIF2-BACCR	Q81Y07 bacillus ce
991	5	1.8	80	1	ACPI-PROMA	Q7V9J0 prochloroco
992	5	1.8	80	1	ACPI-PROMX	Q7V9J6 prochloroco
993	5	1.8	80	1	ACPI-SYNPX	Q7U9V9 synechococc
994	5	1.8	80	1	EXTS-VIBPA	Q87F18 vibrio para
995	5	1.8	80	1	EXTS-VIBBV	Q8FAS5 vibrio vuln
996	5	1.8	80	1	Y359-STREPY	Q9A9B8 streptococc
997	5	1.8	80	1	YF27-METTF	P29563 methanobact
998	5	1.8	80	1	YVDE-VACCV	P04307 vaccinia vi
999	5	1.8	81	1	ACPI-GUITH	P29189 guillardia
1000	5	1.8	81	1	ACPI-RHOBA	Q7UYU2 rhodospirill

ALIGNMENTS

RESULT 1
PLPA_PASHA STANDARD; PRT; 277 AA.

AC Q08868; 007363; 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Outer membrane lipoprotein 1 precursor (PLP1).
GN PLPA.
OS Pasteurella haemolytica.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Mannheimia.
OX NCBI_TaxID=75985;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Serotype A1;
RX MEDLINE=93328110; PubMed=8335249;
RA Murphy G.L., Whitworth L.C.;
RT "Analysis of tandem, multiple genes encoding 30-kDa membrane proteins
in Pasteurella haemolytica A1";
RL Gene 129:107-111(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Serotype A1;
RX MEDLINE=9401378; PubMed=8406866;
RA Cooney B.O., Lo R.Y.C.;
RT "Three contiguous lipoprotein genes in Pasteurella haemolytica A1
which are homologous to a lipoprotein gene in Haemophilus influenzae
type B";
RL Infect. Immun. 61:4682-4688(1993).
CC -1- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
anchor (Probable).
CC -1- SIMILARITY: Belongs to the nlpa lipoprotein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L11037; AAA25538.1; -;
DR EMBL; M91072; AAA25541.1; -;
DR EMBL; L16627; AAA25546.1; -;
DR PIR; JN0751; JN0751.
DR InterPro; IPR004872; Lipoprotein_9.
DR InterPro; IPR000437; ProC_Lipoprot_5.
DR InterPro; IPR004478; YaeC.
DR Pfam; PF03180; Lipoprotein_9; 1.

DR TIGRfam; TIGR00363; TIGR00363; 1.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN_1.
KW Outer membrane; Lipoprotein; Signal; Multigene family; Palmitate.
FT SIGNAL 1
FT CHAIN 20 277
FT LIPID 20 20
FT LIPID 20 20
FT LIPID 20 20
FT CONFLICT 30 30
FT CONFLICT 156 156
FT CONFLICT 160 160
FT CONFLICT 197 200
SQ SEQUENCE 277 AA; 29992 MW; 753B76503026E126 CRC64;
Query Match 12.3%; Score 34; DB 1; Length 277;
Best Local Similarity 100.0%; Pred. No. 9.9e-26;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 140 IAVPNDPSNLTARALILEKQGLIKLKNITLIFST 173
Db 141 IAVNPDSNLTARALILEKQGLIKLKNITLIFST 174

RESULT 2
METQ_HAEIN STANDARD; PRT; 273 AA.

AC P31728;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable D-methionine-binding lipoprotein metQ precursor
(28 kDa outer membrane protein).
GN METQ OR HU2A OR HI0620.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Serotype B;
RX MEDLINE=91100034; PubMed=1987077;
RA Charyangam M., Smith A.L., Moseley S.L., Kuehn M., Jenny P.;
RT "Contribution of a 28-kilodalton membrane protein to the virulence of
Haemophilus influenzae";
RL Infect. Immun. 59:600-608(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shrivley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon K.C.,
RA Fine L.D., Fritchman J.L., Geoghegan N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
Rd";
RL Science 269:496-512(1995).
RN [3]
RP IDENTIFICATION BY MASS SPECTROMETRY.
RX MEDLINE=20137488; PubMed=10675023;
RA Langen H., Takacs B., Myers S., Berndt P., Iahn H.W., Wipf B.,
RA Gray C., Fountoulakis M.;
RT "Two-dimensional map of the proteome of Haemophilus influenzae";
CC Electrophoresis 21:411-429(2000).
CC -1- FUNCTION: This protein is a component of a D-methionine
CC permease, a binding protein-dependent, ATP-driven transport
CC system (By similarity).
CC -1- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
CC anchor (Probable).
CC -1- SIMILARITY: Belongs to the nlpa lipoprotein family.


```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M59804; AAA24939.1; -
DR EMBL; U32744; AAC22279.1; -
DR PIR; B64082; B64082.
DR TIGR; H10620; -
DR InterPro; IPR004872; Lipoprotein_9.
DR InterPro; IPR000437; Prok_lipoprot_S.
DR InterPro; IPR004478; YaeC.
DR Pfam; PF03180; Lipoprotein_9; 1.
DR TIGRfam; TIGR00363; TIGR00363; 1.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
DR Transprot; Amino-acid transport; Outer membrane; Lipoprotein; Signal;
DR Complete proteome; Palmitate.
DR SIGNAL 1 20 PROBABLE.
FT CHAIN 21 273 METO.
FT LIPID 21 21 N-palmitoyl cysteine (Probable).
FT LIPID 21 21 S-diacylglycerol cysteine (Probable).
FT CONFLICT 249 249 I -> V (IN REF. 1).
SQ SEQUENCE 273 AA; 29828 MW; 429838A8AC7DD7D CRC64;

Query Match 6.9%; Score 19; DB 1; Length 273;
Best Local Similarity 100.0%; Pred. No. 6.3e-11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 LNNLVIVGNTFYVPLAGYS 125
DB 104 LNNLVIVGNTFYVPLAGYS 122

RESULT 3
ID PUPC_PASHA STANDARD; PRT; 263 AA.
AC Q08870; Q07365;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Outer membrane lipoprotein 3 precursor (PUP3).
GN PUPC.
OS Pasteurella haemolytica.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Mannheimia.
OX NCBI_TaxID=75985;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=serotype A1;
RC MEDLINE=93338110; PubMed=8335249;
RA Murphy G.L.; Whitworth L.C.;
RT "Analysis of tandem, multiple genes encoding 30-kDa membrane proteins
RT in Pasteurella haemolytica A1."
RL Gene 129:107-111(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=serotype A1;
RC MEDLINE=9401378; PubMed=8406866;
RA Cooney B.J.; Lo R.Y.C.;
RT "Three contiguous lipoprotein genes in Pasteurella haemolytica A1
RT which are homologous to a lipoprotein gene in Haemophilus influenzae
RT type B."
RL Infect. Immun. 61:4682-4688(1993).
CC -1- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
CC anchor (Probable).
CC -1- SIMILARITY: Belongs to the alpha lipoprotein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L11037; AAA25540.1; -
DR EMBL; L16627; AAA25548.1; ALT_INTT.
DR PIR; JN0753; JN0753.
DR InterPro; IPR004872; Lipoprotein_9.
DR InterPro; IPR000437; Prok_lipoprot_S.
DR InterPro; IPR004478; YaeC.
DR Pfam; PF03180; Lipoprotein_9; 1.
DR TIGRfam; TIGR00363; TIGR00363; 1.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
DR Outer membrane; Lipoprotein; Signal; Multigene family; Palmitate.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 263 OUTER MEMBRANE LIPOPROTEIN 3.
FT LIPID 20 20 N-palmitoyl cysteine (Probable).
FT LIPID 20 20 S-diacylglycerol cysteine (Probable).
FT CONFLICT 103 103 T -> S (IN REF. 2).
FT CONFLICT 200 208 ONGLPTKD -> KTVANOS (IN REF. 2).
SQ SEQUENCE 263 AA; 29093 MW; 68FFD8460BD579CB CRC64;

Query Match 4.3%; Score 12; DB 1; Length 263;
Best Local Similarity 100.0%; Pred. No. 0.0005;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 148 NLAARLLLEKQ 159
DB 135 NLAARLLLEKQ 146

RESULT 4
ID METO_PASMU STANDARD; PRT; 276 AA.
AC Q9CK95;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable D-methionine-binding lipoprotein metQ precursor
DE (Outer membrane lipoprotein 1).
GN METO OR PIPA OR PM1730.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Pm70;
RC MEDLINE=21145866; PubMed=11248100;
RA May B.J.; Zhang Q.; Li L.L.; Paustian M.L.; Whittam T.S.; Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70."
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -1- FUNCTION: This protein is a component of a D-methionine
CC permease, a binding protein-dependent, ATP-driven transport
CC system (by similarity).
CC -1- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
CC anchor (Probable).
CC -1- SIMILARITY: Belongs to the alpha lipoprotein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE006210; AAK03814.1; -
DR InterPro; IPR004872; Lipoprotein_9.
DR InterPro; IPR000437; Prok_lipoprot_S.

```

DR InterPro: IPR004478; YaeC.
 DR Pfam: PF03180; Lipoprotein 9; 1.
 DR TIGRfam: TIGR00363; TIGR00363; 1.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 DR Transport; Amino-acid transport; Outer membrane; Lipoprotein; Signal;
 KM Complete proteome; Palmitate.
 FT SIGNAL 1 20 PROBABLE.
 FT CHAIN 21 276 PROBABLE D-METHIONINE-BINDING LIPOPROTEIN
 FT LIPID 21 21 MEMO.
 FT LIPID 21 21 N-palmitoyl cysteine (Probable).
 FT SEQUENCE 276 AA; 30232 MW; 7AFD62A687D624 CRC64;
 SQ
 Query Match 4.3%; Score 12; DB 1; Length 276;
 Best Local Similarity 100.0%; Pred. No. 0.00052;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 155 LLEKOGILKLD 166
 DB 155 LLEKOGILKLD 166
 RESULT 5
 METQ_ECO57 STANDARD; PRT; 271 AA.
 AC Q8X8V9;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE D-methionine-binding lipoprotein metQ precursor.
 GN METQ OR Z0209 OR ECO5195.
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074933; Pubmed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posifal G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimallanta E.T., Potamousis K.,
 RA Aoodaga J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
 RL Nature 409:529-533(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; Pubmed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsuda E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
 RA Kubara S., Shiba T., Hattori M., Shinagawa H.;
 RA "Complete genome sequence of enterohaemorrhagic Escherichia coli
 RT O157:H7 and genomic comparison with a laboratory strain K-12";
 RL DNA Res. 8:11-22(2001).
 CC -1- FUNCTION: This protein is a component of a D-methionine
 CC permease, a binding protein-dependent, ATP-driven transport
 CC system (by similarity).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 CC (Probable).
 CC -1- MISCELLANEOUS: The metNIQ system is also to be able to transport
 CC the toxic methionine analog alpha-methyl-methionine (By
 CC similarity).
 CC -1- SIMILARITY: Belongs to the nlpA lipoprotein family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE005195; AAG54499.1; -
 CC EMBL; AP002550; BAB33622.1; -
 CC PIR; G85504; G85504.
 CC PIR; G90653; G90653.
 CC InterPro; IPR004872; Lipoprotein_9.
 CC InterPro; IPR00437; Prok_Lipoprot_S.
 CC InterPro; IPR004478; YaeC.
 CC Pfam; PF03180; Lipoprotein 9; 1.
 CC TIGRfam; TIGR00363; TIGR00363; 1.
 CC PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 CC Transport; Amino-acid transport; Membrane; Lipoprotein; Signal;
 KM Complete proteome; Palmitate.
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 271 D-METHIONINE-BINDING LIPOPROTEIN METQ.
 FT LIPID 23 23 N-palmitoyl cysteine (Potential).
 FT LIPID 23 23 S-diacylglycerol cysteine (Potential).
 SQ SEQUENCE 271 AA; 29416 MW; 39D7DA1B5CD2B8B CRC64;
 Query Match 4.0%; Score 11; DB 1; Length 271;
 Best Local Similarity 100.0%; Pred. No. 0.0049;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 224 FVEDKSPYVN 234
 DB 219 FVEDKSPYVN 229
 RESULT 6
 METQ_SALTI STANDARD; PRT; 271 AA.
 AC Q8Z9Z2;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE D-methionine-binding lipoprotein metQ precursor.
 GN METQ OR STY0272 OR T0248.
 OS Salmonella typhi.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=601;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CT18;
 RX MEDLINE=21534947; Pubmed=11677608;
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 RA Felwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagsis K.,
 RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
 RA Quail M.A., Rutherford K., Simmonds M., Skelton J., Stevens K.,
 RA Whitehead S., Barrrell B.G.;
 RT "Complete genome sequence of a multiple drug resistant Salmonella
 RT enterica serovar Typhi CT18";
 RL Nature 413:848-852(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Ty2 / ATCC 700931;
 RX MEDLINE=22531367; Pubmed=12644504;
 RA Deng W., Iiou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
 RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
 RA "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
 RT and CT18";
 RL J. Bacteriol. 185:2330-2337(2003).
 CC -1- FUNCTION: This protein is a component of a D-methionine
 CC permease, a binding protein-dependent, ATP-driven transport
 CC system (by similarity).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 CC (Probable).
 CC -1- MISCELLANEOUS: The metNIQ system is also to be able to transport

```
CC the toxic methionine analog alpha-methyl-methionine (By
CC similarity).
CC -1- SIMILARITY: Belongs to the alpha lipoprotein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AL627266; CAD08705.1; -.
CC EMBL: AB016834; AAC67977.1; -.
CC InterPro: IPR004872; Lipoprotein_9.
CC InterPro: IPR000437; Prok_lipoprot_5.
CC InterPro: IPR004478; YaeC.
CC Pfam: PF03180; Lipoprotein_9; 1.
CC TIGRfam: TIGR00363; TIGR00363; 1.
CC PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
CC Transport; Amino-acid transport; Membrane; Lipoprotein; Signal;
CC Complete proteome; Palmitate.
CC SIGNAL
CC CHAIN 1 22
CC FT 23 271 D-METHIONINE-BINDING LIPOPROTEIN METQ.
CC FT 23 23 N-palmitoyl cysteine (Potential).
CC FT 23 23 S-diacylglycerol cysteine (Potential).
CC FT 23 23 S-diacylglycerol cysteine (Potential).
CC SEQUENCE 271 AA; 29465 MW; 49B62C4CF96D9613 CRC64;

Query Match 4.0%; Score 11; DB 1; Length 271;
Best Local Similarity 100.0%; Pred. No. 0.0049;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 224 FVEDKDSPPYV 234
DB 219 FVEDKDSPPYV 229

RESULT 7
METQ_SALTY STANDARD; PRT; 271 AA.
AC 082RNI;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE D-methionine-binding lipoprotein metQ precursor.
GN METQ OR STM0245.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RX MEDLINE=21534948; PubMed=1577609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2."
RL Nature 413:852-856 (2001).
CC -1- FUNCTION: This protein is a component of a D-methionine
CC permease, a binding protein-dependent, ATP-driven transport
CC system (By similarity).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC (Probable).
CC -1- MISCELLANEOUS: The metQ system is also to be able to transport
CC the toxic methionine analog alpha-methyl-methionine (By
CC similarity).
CC -1- SIMILARITY: Belongs to the alpha lipoprotein family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AB008706; AAL19208.1; -.
CC STyGene: SG7272?; metQ.
CC InterPro: IPR004872; Lipoprotein_9.
CC InterPro: IPR000437; Prok_lipoprot_5.
CC InterPro: IPR004478; YaeC.
CC Pfam: PF03180; Lipoprotein_9; 1.
CC TIGRfam: TIGR00363; TIGR00363; 1.
CC PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
CC Transport; Amino-acid transport; Membrane; Lipoprotein; Signal;
CC Complete proteome; Palmitate.
CC SIGNAL
CC CHAIN 1 22
CC FT 23 271 D-METHIONINE-BINDING LIPOPROTEIN METQ.
CC FT 23 23 N-palmitoyl cysteine (Potential).
CC FT 23 23 S-diacylglycerol cysteine (Potential).
CC FT 23 23 S-diacylglycerol cysteine (Potential).
CC SEQUENCE 271 AA; 29435 MW; 3FED41E14FBD707 CRC64;

Query Match 4.0%; Score 11; DB 1; Length 271;
Best Local Similarity 100.0%; Pred. No. 0.0049;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 224 FVEDKDSPPYV 234
DB 219 FVEDKDSPPYV 229

RESULT 8
RPA9_YEAST STANDARD; PRT; 125 AA.
AC P32529;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE DNA-directed RNA polymerase I 13.7 kDa polypeptide (EC 2.7.7.6)
DE (A12.2).
GN RPA12 OR RRM4 OR YTR063M OR J1747.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RX MEDLINE=93109294; PubMed=8417319;
RA Nogai Y., Yano R., Dodd J., Carles C., Nomura M.;
RT "Gene RRM4 in Saccharomyces cerevisiae encodes the A12.2 subunit of
RT RNA polymerase I and is essential only at high temperatures."
RL Mol. Cell. Biol. 13:114-122 (1993).
RN [1]
RN SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RP MEDLINE=93109294; PubMed=8417319;
RX MEDLINE=52886;
RA STRAIN=52886;
RA Huang M.-E., Manns V., Chut J.-C., Gilbert F.;
RT "Analysis of a 62 kb DNA sequence of chromosome X reveals 36 open
RT reading frames and a gene cluster with a counterpart on chromosome
RT XI."
RL Yeast 12:863-875 (1996).
CC -1- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription
CC of DNA into RNA using the four ribonucleoside triphosphates as
CC substrates.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA) (N).
CC -1- SUBUNIT: RNA polymerase I consists of 14 different subunits.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- MISCELLANEOUS: Three distinct zinc-containing RNA polymerases are
CC found in eukaryotic nuclei: polymerase I for the ribosomal RNA
CC precursor, polymerase II for the mRNA precursor, and polymerase
```

-1- SIMILARITY: Belongs to the archaeal rpm / eukaryotic RPA12/RPB9/RPCL RNA polymerase family.
 -1- SIMILARITY: Contains 1 ribbon-type zinc finger.

 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to [license@isb-sid.ch](mailto:license@isb-sib.ch)).

EMBL; L00708; AAA3492.1; -
 EMBL; L35564; AAB59319.1; -
 EMBL; Z49563; CAA89591.1; -
 EMBL; L47993; AAB39289.1; -
 PIR; A48107; A48107.
 DR HSP; Q56254; 1QY.
 DR Germonline; 141896; -
 DR SGD; S0003824; RPA12.
 DR InterPro; IPR001529; RNA POL.
 DR InterPro; IPR001222; TFIIS.
 DR Pfam; PF02150; RNA_POL_M_15KD; 1.
 DR Pfam; PF01096; TFIIS; 1.
 DR SMART; SM00661; RPO19; 1.
 DR SMART; SM00440; ZNF_C2C2; 1.
 DR PROSITE; PS00466; TFIIS; 1.
 DR PROSITE; PS01030; RNA_POL_M_15KD; 1.
 KM Transferase; DNA-directed RNA polymerase; Transcription;
 KM Nuclear protein; zinc-finger.
 FT ZN_FING 10 33 ZN-RIBBON (POTENTIAL).
 FT ZN_FING 86 117 ZN-RIBBON (POTENTIAL).
 SQ SEQUENCE 125 AA; 13660 MW; D79372070819987C CRC64;

Query Match 3.3%; Score 9; DB 1; Length 125;
 Best Local Similarity 100.0%; Pred. No. 0.24;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 NEUKDGTI 140
 Db 74 NEUKDGTI 82

RESULT 9
 METQ_VIBCH STANDARD; PRT; 269 AA.
 ID METQ_VIBCH
 AC Q9KJT7;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Probable D-methionine-binding lipoprotein meto precursor.
 GN METQ OR VC0905.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 OX NCBI_Taxid=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=El Tor N16961 / Serotype O1;
 RA MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umeyam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L., Ustachek T., Fleischmann R.D., Newman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RA "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 RT cholerae.";
 RL Nature 406:477-483 (2000).
 CC -1- FUNCTION: This protein is a component of a D-methionine
 CC pemase: a binding protein-dependent, ATP-driven transport
 CC system (by similarity).

-1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 (Probable).
 -1- SIMILARITY: Belongs to the alpha lipoprotein family.

 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to [license@isb-sid.ch](mailto:license@isb-sib.ch)).

EMBL; AE004174; AAF94067.1; ALT_INIT.
 DR TIGR; VC0905; -
 DR InterPro; IPR004872; Lipoprotein_9.
 DR InterPro; IPR000437; Prok_Lipoprot_S.
 DR InterPro; IPR004478; YaeC.
 DR Pfam; PF01180; Lipoprotein_9; 1.
 DR TIGRFAMs; TIGR00363; TIGR00363; 1.
 DR PROSITE; PS00013; PROKR_LIPOPROTEIN; 1.
 KM Transport; Amino-acid transport; Membrane; Lipoprotein; Signal;
 KM Complete proteome; Palmitate.
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 269 PROBABLE D-METHIONINE-BINDING LIPOPROTEIN
 FT METQ.
 FT LIPID 23 23 N-palmitoyl cysteine (Potential).
 FT LIPID 23 23 S-diacylglycerol cysteine (Potential).
 SQ SEQUENCE 269 AA; 29067 MW; 39E2570E3F184D6 CRC64;

Query Match 3.3%; Score 9; DB 1; Length 269;
 Best Local Similarity 100.0%; Pred. No. 0.46;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 112 IVGNTFYYP 120
 Db 105 IVGNTFYYP 113

RESULT 10
 MCB_HUMAN STANDARD; PRT; 391 AA.
 ID MCB_HUMAN
 AC P04220;
 DT 20-MAR-1987 (Rel. 04, Created)
 DT 20-MAR-1987 (Rel. 04, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Ig mu heavy chain disease protein (BOR).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE.
 RC MEDLINE=84184186; PubMed=6425189;
 RA Barnikol-Watanabe S., Mhaesco E., Mhaesco C., Barnikol H.U.,
 RA Hilschmann N.;
 RA "The primary structure of mu-chain-disease protein BOR. Peculiar
 RT amino-acid sequence of the N-terminal 42 positions."
 RL Hoppe-Seyler's Z. Physiol. Chem. 365:105-118 (1984).
 CC -1- MISCELLANEOUS: This protein has no V region homology or CH1
 CC region.
 CC PIR; A02163; MHHTBT.
 DR HSP; F01857; 1FCT1.
 DR GO; GO:0005624; C:membrane fraction; NAS.
 DR GO; GO:0003823; P:antigen binding; TAS.
 DR GO; GO:0006955; P:immune response; NAS.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003597; IG cl.
 DR InterPro; IPR003006; IG_MHC.
 DR Pfam; PF00477; Ig; 3.
 DR SMART; SM00407; IGC1; 2.
 DR PROSITE; PS00835; IG_LIKE; 3.
 DR PROSITE; PS00290; IG_MHC; 3.
 KM Immunoglobulin domain; Immunoglobulin C region.

FT DOMAIN 1 42 PRE-C-PART (NO V REGION HOMOMOLOGY).
 FT DOMAIN 43 155 CH2.
 FT DOMAIN 156 261 CH3.
 FT DOMAIN 262 391 CH4.
 SQ SEQUENCE 391 AA; 43057 MW; 9100843AFC021A CRC64;

Query Match 3.3%; Score 9; DB 1; Length 391;
 Best Local Similarity 100.0%; Pred. No. 0.64;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 TDEVEARAK 265
 DB 99 TDEVEARAK 107

RESULT 11
 YD97_CAMTE STANDARD; PRT; 74 AA.
 AC 09PMEO;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein Cj1397.
 GN Cj1397.
 OS Campylobacter jejuni.
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 OC Campylobacteraceae; Campylobacter.
 OX NCBI_TaxID=197;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCCTC 11168;
 RX MEDLINE=20150912; PubMed=10688204;
 RA Parkhill J., Wren B.W., Mungall K., Kelsey J.M., Churcher C.,
 RA Basham D., Chillingworth T., Davies R.M., Feltham T., Holtroyd S.,
 RA Jorgensen K., Kariyasek A.V., Moule S., Pallen M.J., Penn C.W.,
 RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
 RA Whitehead S., Barrett B.G.;
 RA "The genome sequence of the food-borne pathogen Campylobacter jejuni
 RT reveals hypervariable sequences.";
 RL Nature 403:665-668(2000).
 RT -1- SIMILARITY: TO H. PYLORI HP0585.1/JHP0533.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; AJ139078; CAB73821.1; -.
 DR PIR; E81284; E81284.
 DR InterPro; IPR007167; FeoA.
 DR Pfam; PF04023; FeoA; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 74 AA; 8363 MW; E81BC8A8369DDC4 CRC64;

Query Match 2.9%; Score 8; DB 1; Length 74;
 Best Local Similarity 100.0%; Pred. No. 1.5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 TUNELKDG 137
 DB 2 TUNELKDG 9

RESULT 12
 METQ_ECOLI STANDARD; PRT; 271 AA.
 AC P28655;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-UTL-1993 (Rel. 26, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE D-methionine-binding lipoprotein metQ precursor.
 GN METQ OR B0197.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RA Miyamoto K.;
 RL Submitted (APR-1993) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blatter F.R., Plunkett G., Illi, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Valdes J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RL "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / W3110;
 RA Takemoto K., Mori H., Murayama N., Kataoka K., Yano M., Itoh T.,
 RA Yamamoto Y., Inokuchi H., Miki T., Hatada E., Fukuda R., Ichihara S.,
 RA Mizuno T., Makino K., Nakata A., Yura T., Samei G., Mizobuchi K.;
 RL "Systematic sequencing of the Escherichia coli genome: analysis of the
 RL 4.0 - 6.0 min (189,987 - 281,416bp) region.";
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RA Schramm S., Duncan M., Allen E., Araujo R., Aparicio A., Chung E.,
 RA Davis K., Federspiel N., Hyman R., Kalman S., Kong C., Kurd O.,
 RA Lahtkari D., Lew H., Lin D., Namath A., Oefner P., Roberts D.,
 RA Davis R.W.;
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 127-271 FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=93094132; PubMed=1459951;
 RA Gervais F.G., Drapeau G.R.;
 RL "Identification, cloning, and characterization of rcsF, a new
 RL regulator gene for exopolysaccharide synthesis that suppresses the
 RL division mutation fts284 in Escherichia coli K-12.";
 RL J. Bacteriol. 174:8016-8022(1992).
 RN [6]
 RP FUNCTION.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=22159950; PubMed=12169620;
 RA Gal J., Szvechnik A., Schmeil R., Kalman M.;
 RL "The metQ D-methionine transporter locus of Escherichia coli is an ABC
 RL transporter gene cluster.";
 RL J. Bacteriol. 184:4930-4932(2002).
 CC -1- FUNCTION: This protein is a component of a D-methionine
 CC permease, a binding protein-dependent, ATP-driven transport
 CC system.
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 CC (Probable).
 CC -1- MISCELLANEOUS: The metQ system is also to be able to transport
 CC the toxic methionine analog alpha-methyl-methionine.
 CC -1- SIMILARITY: Belongs to the alpha lipoprotein family.
 CC -1- CAUTION: REF.S SEQUENCE DIFFERS FROM THAT SHOWN IN POSITION 217
 CC ONWARD DUE TO A FRAMESHIFT.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC

DR EMBL; D15061; BAA03657.1; -
 DR EMBL; AE000129; AAC73308.1; -
 DR EMBL; D83536; BAA78784.1; -
 DR EMBL; U70214; AAB08625.1; -
 DR EMBL; U04474; AAA24507.1; ALT_FRAME.
 DR PIR; E64744; E64744.
 DR SWISS-2DPAGE; P28635; COLI.
 DR Ecogen; E61504; metC.
 DR InterPro; IPR004872; Lipoprotein_9.
 DR InterPro; IPR000437; Prok_lipoprot_5.
 DR InterPro; IPR004478; YaeC.
 DR Pfam; PF0180; Lipoprotein 9; 1.
 DR TIGRFAMs; TIGR00363; TIGR00363; 1.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 DR Transport; Amino-acid transport; Membrane; Lipoprotein; Signal;
 KW Complete proteome; Palmitate.
 FT SIGNAL 1 22
 FT CHAIN 23 271
 FT LIPID 23 23
 FT LIPID 23 23
 FT CONFID 200 200
 FT SEQUENCE 271 AA; 29431 MW; B50C6C6B5CD2BF7 CRC64;
 Query Match 2.9%; Score 8; DB 1; Length 271;
 Best Local Similarity 100.0%; Pred. No. 4.5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 113 VGNTPVFP 120
 DB 106 VGNTPVFP 113

RESULT 13
 SM4G_MOUSE STANDARD; PRT; 837 AA.
 AC Q9WU7; 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Semaphorin 4G precursor.
 GN SEMA4G.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RX MEDLINE=99425180; Pubmed=10495281;
 RA Li H., Wu D.K., Sullivan S.L.;
 RT "Characterization and expression of sema4g, a novel member of the
 semaphorin gene family.";
 RL Mech. Dev. 87:169-173(1999).
 CC -1- FUNCTION: MAY PLAY A ROLE IN AXON GUIDANCE.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: BRAIN, SPINAL CORD, AND SEVERAL SENSORY ORGANS
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED EARLY IN DEVELOPMENT.
 CC -1- SIMILARITY: Belongs to the semaphorin family.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL; AF134918; AAB0541.1; -
 DR MGD; MGI:1347047; Sema4g.
 DR InterPro; IPR007110; Ig-like.

DR InterPro; IPR003599; Ig.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR002165; Plexin_repeat.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF01437; PSI; 1.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SMC0409; IG; 1.
 DR SMART; SMC0423; PSI; 1.
 DR SMART; SMC0630; Sema; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 DR Signal; transmembrane; Immunoglobulin domain; Multigene family;
 KW Neurogenesis; Developmental protein; Glycoprotein.
 FT SIGNAL 1 17
 FT CHAIN 18 837
 FT DOMAIN 18 673
 FT TRANSMEM 674 694
 FT DOMAIN 695 837
 FT DOMAIN 56 487
 FT DOMAIN 505 556
 FT DOMAIN 565 647
 FT DOMAIN 563 566
 FT DOMAIN 762 773
 FT DISULFD 582 630
 FT CARBOHYD 55 55
 FT CARBOHYD 111 111
 FT CARBOHYD 126 126
 FT CARBOHYD 386 386
 FT CARBOHYD 540 540
 FT CARBOHYD 596 596
 FT SEQUENCE 837 AA; 92378 MW; SC69C9C8A545E31 CRC64;
 Query Match 2.9%; Score 8; DB 1; Length 837;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 146 PSNLARAL 153
 DB 585 PSNLARAL 592

RESULT 14
 SM4G_HUMAN STANDARD; PRT; 838 AA.
 AC Q9WU9; 09DEC3;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Semaphorin 4G precursor.
 GN SEMA4G OR KIAA1619.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RN RP SEQUENCE FROM N.A.
 RC Heath P.;
 RA Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=20450683; Pubmed=10997877;
 RX Nagase T., Kikuno R., Nakayama M., Hirosewa M., Ohara O.;
 RA "Prediction of the coding sequences of unidentified human genes.
 XVIII. The complete sequences of 100 new cDNA clones from brain which
 code for large proteins in vitro.";
 RL DNA Res. 7:273-281(2000).
 CC -1- FUNCTION: MAY PLAY A ROLE IN AXON GUIDANCE (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: Belongs to the semaphorin family.
 CC -1- SIMILARITY: Contains 1 Sema domain.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; AL133215; -; NOT ANNOTATED CDS.
 CC EMBL; AB046839; BAB13445.1; ALT_INIT.
 CC Genew; HGNCL10735; SEMA4G.
 CC InterPro; IPR007110; IG-like.
 CC InterPro; IPR003599; IG.
 CC InterPro; IPR003659; Plexin-like.
 CC InterPro; IPR002165; Plexin_repeat.
 CC InterPro; IPR001627; Sema.
 CC Pfam; PF01437; PSI; 1.
 CC Pfam; PF01403; Sema; 1.
 CC SMART; SM00409; IG; 1.
 CC SMART; SM00423; PSI; 1.
 CC SMART; SM00630; Sema; 1.
 CC PROSITE; PS50835; IG_Like; 1.
 CC Signal; Transmembrane; Immunoglobulin domain; Multigene family;
 CC Neurogenesis; Developmental protein; Glycoprotein.
 CC SIGNL 1
 CC CHAIN 17
 CC FT 18 838 SEMAPHORIN 4G. (POTENTIAL).
 CC FT DOMAIN 18 675 EXTRACELLULAR (POTENTIAL).
 CC FT TRANSMEM 676 696 POTENTIAL.
 CC FT DOMAIN 697 838 CYTOPLASMIC (POTENTIAL).
 CC FT DOMAIN 56 489 SEMA.
 CC FT DOMAIN 507 558 PSI.
 CC FT DOMAIN 567 649 IG-LIKE C2-TYPE.
 CC FT DOMAIN 565 568 POLY-PRO.
 CC FT DOMAIN 763 774 POLY-PRO.
 CC FT DISULFID 584 632 BY SIMILARITY.
 CC FT CARBOHYD 111 111 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC FT CARBOHYD 126 126 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC FT CARBOHYD 388 388 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC FT CARBOHYD 542 542 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC FT CARBOHYD 598 598 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC FT CONFLICT 543 543 R -> RSQGR (IN REF. 2).
 CC SQ SEQUENCE 838 AA; 91496 MW; 9B281AE8681F245 CRC64;

Query Match 2.9%; Score 8; DB 1; Length 88;
 Best Local Similarity 100.0%; Pred.No. 12;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 PSN1ARAL 153
 DB 587 PSN1ARAL 594

RESULT 15
 ID LBN MOUSE STANDARD; PRT; 1220 AA.
 AC OIKG2; Q8BRF3;
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Limbin.
 GN EVC2 OR LBN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RC TISSUE=Limb bud;
 RX MEDLINE=2155879; Pubmed=12136126;
 RA Takea H., Takami M., Oguni T., Tsuji T., Yoneda K., Sato H.,
 RA Ihara N., Itoh T., Kata S.R., Mishina Y., Womack J.E., Moritomo Y.,
 RA Sugimoto Y., Kunieda T.,
 RA "Positional cloning of the gene limbin responsible for bovine
 RT chondrodysplastic dwarfism";

RL Proc. Natl. Acad. Sci. U.S.A. 99:10549-10554 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Kidney;
 RX MEDLINE=22388257; Pubmed=12477932;
 RA Stussberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shamen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diachenko L., Martina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.J., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loguettano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Holyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Small D.E.,
 RA Scherch A., Schein J.E., Jones S.J.W., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences";
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 [3]
 RN SEQUENCE OF 21-744 FROM N.A.
 RC STRAIN=C57BL/6J;
 RX MEDLINE=22354683; Pubmed=12466851;
 RA Okazaki Y., Futuro N., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nixaido I., Osato N., Saito R., Suzuki H., Yamana I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schobach C., Gajdoori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Brad D., Brusci V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalia E., Dragan T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Kongsawa A., Kurochkin I.V., Lee Y., Lemhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Nimata K., Okido T., Pavan W.J., Petrea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius U.V., Qi D., Ramchandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Sempke C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wainwright C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wyszewski-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Maki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.,
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs";
 Nature 420:563-573 (2002).
 RL -1- FUNCTION: Plays a critical role in bone formation and skeletal
 development. May be involved in early embryonic morphogenesis.
 CC -1- TISSUE SPECIFICITY: Expressed in long and cranial bones, kidney
 CC and heart. Strongly expressed in proliferating chondrocytes,
 CC osteoblasts and osteoclasts.
 CC -1- DEVELOPMENTAL STAGE: Found in the embryo at day E7, E11, E15, and
 CC E17. At the limb bud formation stage E11, it is expressed in fore-
 CC and hindlimb buds, branchial arches, and facial primordia.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@sb-sib.ch.
CC -----
DR EMBL; AB083066; BAC06589.1; -
DR EMBL; BC037473; AAC37473.1; -
DR EMBL; AK044977; BAC32167.1; ALT_INIT.
DR MGD; MGI:191575; Evc2.
KM Transmembrane; Colled coil.
FT DOMAIN 211 231 POTENTIAL.
FT DOMAIN 355 404 COILED COIL (POTENTIAL).
FT DOMAIN 563 644 COILED COIL (POTENTIAL).
FT DOMAIN 854 875 COILED COIL (POTENTIAL).
FT DOMAIN 920 1005 COILED COIL (POTENTIAL).
SQ SEQUENCE 1220 AA; 137638 MW; E67671714A866BD CRC64;

Query Match 2.9%; Score 8; DB 1; Length 1220;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 80 SAVSKGEL 87
Db 457 SAVSKGEL 464

RESULT 16
ID HFB_XANAC STANDARD; PRT; 103 AA.
AC O8P878;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integration host factor beta-subunit (IHf-beta).
GN IHFB OR HIMD OR XAC2297.
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., de Amaral A.M., Bertolini M.C., Canargo L.E.A.,
RA Camarotte G., Canavan F., Cardoso J., Chamberggo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A.F., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spindola L.A.F., Takita M.A., Tamura R.S., Teixeira E.C., Tezza R.I.D.,
RA Tirindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities";
RL Nature 417:459-463(2002).
CC -1- FUNCTION: This protein is one of the two subunits of integration
CC host factor, a specific DNA-binding protein that functions in
CC genetic recombination as well as in transcriptional and
CC transnational control (By similarity).
CC -1- SUBUNIT: Heterodimer of an alpha and a beta chain (By similarity).
CC -1- SIMILARITY: Belongs to the bacterial histone-like protein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sb-sib.ch).
CC -----

DR EMBL; AE018666; AAM37150.1; -
DR HAVAP; MF_00381; -; 1.
DR InterPro; IPR000119; Bac_DNAbind.
DR InterPro; IPR005685; IHf_beta.
DR Pfam; PF00216; Bac_DNA_binding; 1.
DR PRINTS; PRO1727; DNABINDINGHU.
DR ProDom; PD000945; Bac_DNABIND; 1.
DR TIGRfam; TIGR00989; hfp; 1.
DR PROSITE; PS00045; HISTONE_LIKE; 1.
KM DNA-binding; transcription regulation; DNA recombination;
KM translation regulation; Complete proteome.
SQ SEQUENCE 103 AA; 11373 MW; 8D5F59687D3FDB6 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 103;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 199 DDVDLAV 205
Db 20 DDVDLAV 26

RESULT 17
ID HFB_XANCP STANDARD; PRT; 103 AA.
AC O8P8P6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integration host factor beta-subunit (IHf-beta).
GN IHFB OR HIMD OR XCC2193.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A. / NCPPB 528;
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., de Amaral A.M., Bertolini M.C., Canargo L.E.A.,
RA Camarotte G., Canavan F., Cardoso J., Chamberggo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A.F., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spindola L.A.F., Takita M.A., Tamura R.S., Teixeira E.C., Tezza R.I.D.,
RA Tirindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities";
RL Nature 417:459-463(2002).
CC -1- FUNCTION: This protein is one of the two subunits of integration
CC host factor, a specific DNA-binding protein that functions in
CC genetic recombination as well as in transcriptional and
CC transnational control (By similarity).
CC -1- SUBUNIT: Heterodimer of an alpha and a beta chain (By similarity).
CC -1- SIMILARITY: Belongs to the bacterial histone-like protein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sb-sib.ch).
CC -----
CC EMBL; AE012326; AAM41473.1; -

```

DR HAMAP: MF_00381; -; 1.
DR InterPro: IPR000119; Bac_DNAbind.
DR Pfam: PF00216; Bac_DNA_binding; 1.
DR PRINTS: PR01727; DNABINDINGH.
DR PRODOM: PD000945; Bac_DNAbind; 1.
DR TIGRFAMs: TIGR00988; hlp; 1.
DR PROSITE: PS00045; HISTONE_LIKE; 1.
KM DNA-binding; Transcription regulation; DNA recombination;
KW Translation regulation; Complete proteome.
SQ SEQUENCE 103 AA; 11403 MW; 8D5F45186237ADB6 CRC64;

Query Match      2.5%; Score 7; DB 1; Length 103;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 DDVDLAV 205
DB 20 DDVDLAV 26

RESULT 18
ID IHFB_XYLFA STANDARD; PRT; 104 AA.
AC QSPAC8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integration host factor beta-subunit (IHFB-beta).
GN IHFB OR HIMD OR XP2437.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=945C; PubMed=10910347;
RX MEDLINE=20365717;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvares A.J.G., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Brites M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrier H.,
RA Colauto N.B., Colombo C., Costa F.R., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.B., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Faciniani A.P., Ferreira A.J.S., Ferreira V.C.A., Fierro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Kieger J.E., Kuramae E.E., Laigret F., Lambis M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Martino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Mizoca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A.J., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A.Jr., Pequeiro J.B.,
RA Queiroz R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E., Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silva J.F., Silveira M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuchioka M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Zechin J., Zetser J.C.,
RT "The genome sequence of the plant pathogen Xylella fastidiosa."
RL Nature 406:151-159(2000).
CC -1- FUNCTION: This protein is one of the two subunits of integration
CC host factor, a specific DNA-binding protein that functions in
CC genetic recombination as well as in transcriptional and
CC translational control (By similarity).
CC -1- SUBUNIT: Heterodimer of an alpha and a beta chain.
CC -1- SIMILARITY: Belongs to the bacterial histone-like protein family.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE004052; AAF85236.1; AUT_INIT.
DR HSSP: P36206; 1982.
DR HAMAP: MF_00381; -; 1.
DR InterPro: IPR000119; Bac_DNAbind.
DR Pfam: PF00216; Bac_DNA_binding; 1.
DR PRINTS: PR01727; DNABINDINGH.
DR PRODOM: PD000945; Bac_DNAbind; 1.
DR SMART: SM00411; BHL; 1.
DR TIGRFAMs: TIGR00988; hlp; 1.
DR PROSITE: PS00045; HISTONE_LIKE; 1.
KM DNA-binding; Transcription regulation; DNA recombination;
KW Translation regulation; Complete proteome.
SQ SEQUENCE 104 AA; 11463 MW; 779514E91FA731CC CRC64;

Query Match      2.5%; Score 7; DB 1; Length 104;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 DDVDLAV 205
DB 20 DDVDLAV 26

RESULT 19
ID IHFB_XYLFT STANDARD; PRT; 104 AA.
AC Q87B8;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Integration host factor beta-subunit (IHFB-beta).
GN IHFB OR HIMD OR PD1455.
OS Xylella fastidiosa (strain Temecula / ATCC 700964).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=183190;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22421331; PubMed=12533478;
RA Van Sluys M.A., de Oliveira M.C., Monteiro-Vitorello C.B.,
RA Miyaki C.Y., Furlan L.R., Camargo L.E.A., da Silva A.C.R., Moon D.H.,
RA Takita W.A., Lemos E.G.M., Machado M.A., Fierro M.I.T., da Silva F.R.,
RA Goldman M.H.S., Goldman G.H., Lemos M.V.F., El-Dorri H., Tsai S.M.,
RA Carrier H., Carraro D.M., de Oliveira R.C., Nunes L.R., Siqueira W.J.,
RA Coutinho L.L., Kimura E.T., Petro E.S., Harakava R., Kuramae E.E.,
RA Martino C.L., Gigliotti E., Abreu I.L., Alves L.M.C., do Amaral A.V.,
RA Bata G.S., Bianco S.R., Brito M.S., Cannavan F.S., Celestino A.V.,
RA da Cunha A.F., Fenille R.C., Fierro J.A., Formighieri E.F., Kishi L.T.,
RA Leoni S.G., Oliveira A.R., Rosa V.E. Jr., Sasaki F.T., Sena U.A.D.,
RA de Souza A.A., Truffi D., Tsuchioka M.H., Yanai G.M., Zatz M.,
RA Civerolo E.L., Simpson A.J.G., Almeida N.F. Jr., Setubal J.C.,
RA Kitajima J.P.;
RT "Comparative analyses of the complete genome sequences of Pierce's
RT disease and citrus variegated chlorosis strains of Xylella
RT fastidiosa."
RL J. Bacteriol. 185:1018-1026(2003).
CC -1- FUNCTION: This protein is one of the two subunits of integration
CC host factor, a specific DNA-binding protein that functions in
CC genetic recombination as well as in transcriptional and
CC translational control (By similarity).
CC -1- SUBUNIT: Heterodimer of an alpha and a beta chain (By similarity).
CC -1- SIMILARITY: Belongs to the bacterial histone-like protein family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; A012558; AAC02929.1; -
 DR HAMAP; MF 00381; -; 1.
 DR InterPro; IPR000119; Bac DNAbind.
 DR Pfam; PF00216; Bac DNA binding; 1.
 DR PRINTS; PRO1727; DNABINDINGHU.
 DR PROSITE; PS00045; HISTONE_LIKE; 1.
 KW DNA-binding; Transcription regulation; DNA recombination;
 KW translation regulation; Complete proteome.
 SQ SEQUENCE 104 AA; 11377 MW; 999677891CC451C9 CRC64;
 Query Match 2.5%; Score 7; DB 1; Length 104;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 199 DDVDLAV 205
 Db 20 DDVDLAV 26
 RESULT 20
 PSP1_PSP1C STANDARD; PRT; 141 AA.
 ID PSP1_PSP1C
 AC 061704;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Plasmacyte-spreading peptide precursor.
 GN PSP1.
 OS Pseudoplusia includens (Soybean looper).
 OC Neoptera; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Eukaryota; Metazoa; Lepidoptera; Glossata; Ditrysia; Noctuoidea;
 OC Noctuidae; Plusiinae; Pseudoplusia.
 CC NCBI_TaxID=76492;
 [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=98440825; PubMed=9753657;
 RX Clark K.D., Witherell A., Strand M.R.;
 RA "Plasmacyte spreading peptide is encoded by an mRNA differentially
 RT expressed in tissues of the moth Pseudoplusia includens.";
 RL Biochem. Biophys. Res. Commun. 250:479-485(1998).
 [2]
 RN SEQUENCE OF 119-141.
 RP TISSUE=Hemolymph;
 RC MEDLINE=97435321; PubMed=9287360;
 RA Clark K.D., Pech L.L., Strand M.R.;
 RT "Isolation and identification of a plasmacyte-spreading peptide
 RT from the hemolymph of the lepidopteran insect Pseudoplusia
 RT includens.";
 RL J. Biol. Chem. 272:23440-23447(1997).
 [3]
 RN STRUCTURE BY NMR OF 119-141.
 RP TISSUE=Hemolymph;
 RC MEDLINE=9988679;
 RX Volkman B.F., Anderson M.E., Clark K.D., Hayakawa Y., Strand M.R.,
 RA Markley J.L.;
 RT "Structure of the insect cytokine peptide plasmacyte-spreading
 RT peptide 1 from Pseudoplusia includens.";
 RL J. Biol. Chem. 274:4493-4496(1999).
 CC -1- FUNCTION: Mediates the spreading of plasmacytes to foreign
 CC surfaces. Plasmacytes are a class of hemocytes involved in insect
 CC cellular immunity.
 CC -1- SIMILARITY: Belongs to the GBP / PSP1 / paralytic peptide family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF062489; AAC16546.1; -
 DR PIR; J0359; J0359.
 DR PDB; 1B1V; 02-DRC-98.
 DR PDB; 1B5N; 13-JAN-99.
 DR InterPro; IPR003463; GBP_PSP.
 DR Pfam; PF02425; GBP_PSP; 1.
 DR Prodom; PD06507; GBP_PSP; 1.
 KW Hemolymph; Signal; 3D-structure.
 FT SIGNAL 1 22
 FT PROPEP 23 118
 FT PEPTIDE 119 141
 FT DISULFID 125 137
 FT STRAND 130 131
 FT STRAND 137 138
 SQ SEQUENCE 141 AA; 14953 MW; 929B4BD55CP21A4 CRC64;
 Query Match 2.5%; Score 7; DB 1; Length 141;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 223 VFVEDKD 229
 Db 65 VFVEDKD 71
 RESULT 21
 ID COAT_ACT1A STANDARD; PRT; 193 AA.
 AC P54890;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Coat protein.
 OS Apple chlorotic leaf spot virus (isolate apple) (ACUSV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flexiviridae;
 OC Trichovirus.
 CC NCBI_TaxID=73472;
 [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=93389448; PubMed=8376968;
 RX Sato K., Yoshikawa N., Takahashi T.;
 RA "Complete nucleotide sequence of the genome of an apple isolate of
 RT apple chlorotic leaf spot virus.";
 RL J. Gen. Virol. 74:1927-1931(1993).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D14996; BAA03643.1; -
 DR PIR; J02185; J02185.
 DR InterPro; IPR008879; Tricho_coat.
 DR Pfam; PF05892; Tricho_coat; 1.
 KW Coat protein.
 SQ SEQUENCE 193 AA; 21395 MW; C81EG30BR33A7COA CRC64;
 Query Match 2.5%; Score 7; DB 1; Length 193;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 168 TNLPEST 174
 Db 126 TNLPEST 132

```

RESULT 22
MRKB_KLEPN STANDARD; PRT; 233 AA.
ID MRKB_KLEPN
AC P21646;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 15-JUN-1999 (Rel. 38, Last annotation update)
DE Chapterone protein mrkb precursor.
MRKB.
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Klebsiella.
NCBI_TaxID=573;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=1A565;
RX MEDLINE=91100388; PubMed=1670938;
RA Allen B.L., Gerlach G.-F., Clegg S.;
RT "Nucleotide sequence and functions of mrk determinants necessary for
RT expression of type 3 fimbriae in Klebsiella pneumoniae.";
RL J. Bacteriol. 173:916-920(1991)
CC -1- FUNCTION: MEDIANES ASSEMBLY OF PILI BY FORMING SOLUBLE MULTIMERIC
CC COMPLEXES WITH PILI SUBUNITS AS AN INTERMEDIATE STEP IN THE
CC ASSEMBLY PROCESS. THIS PROTEIN IS INVOLVED IN TYPE 3 PILI
CC ASSEMBLY.
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- SIMILARITY: Belongs to the periplasmic pilus chapterone family.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M55912; AAA25094.1; -.
DR PIR; C39142; C39142.
DR HSSP; P31697; 1BF8.
DR InterPro: IPR008962; PapD-like.
DR InterPro: IPR001829; Pili chapterone.
DR Pfam; PF00345; Pili_assembly; 1.
DR PRINTS; PR00969; CHAPERONPILI.
DR ProDom; PD001447; Pili chapterone; 1.
DR PROSITE; PS00635; Pili chapterone; 1.
KW Chapterone; Fimbria; Periplasmic; Signal; Immunoglobulin domain.
FT SIGNAL 1
FT CHAIN 19 233 CHAPERONE PROTEIN MRKB.
SQ SEQUENCE 233 AA; 25143 MW; F71EA950AF2317C CRC64;

Query Match 2.5%; Score 7; DB 1; Length 233;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 SAVSKG3 86
Db 206 SAVSKG3 212

```

```

NCBI_TaxID=2097;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=96026346; PubMed=7569993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischman R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.D.,
RA Nguyen D.T., Utterback T.R., Saudak D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Botz K.F., Hu P.-C., Lueder J.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT "The minimal gene complement of Mycoplasma genitalium.";
RL Science 270:397-403(1995).
CC -1- FUNCTION: Essential for recycling GMP and indirectly, cGMP.
CC -1- CATALYTIC ACTIVITY: ATP + GMP = ADP + GDP.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC -1- SIMILARITY: Belongs to the guanylate kinase family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; U39691; AAC71325.1; ALT_INIT.
DR HSSP; P15454; 1GKY.
DR TIGR; MG107; -.
DR HAMAP; MF_00328; -.
DR InterPro: IPR008144; Guanylate_kin.
DR InterPro: IPR008145; Guanylt/Ca.
DR Pfam; PF00625; Guanylate_kin; 1.
DR SMART; SMC0072; GuKc; 1.
DR PROSITE; PS00856; GUANYLATE_KINASE_1; 1.
DR PROSITE; PS50052; GUANYLATE_KINASE_2; 1.
KW Transferase; Kinase; ATP-binding; Complete proteome.
FT NP BIND 63
FT SIGNAL 70
SQ SEQUENCE 240 AA; 27520 MW; 0664E11FE90935D CRC64;

Query Match 2.5%; Score 7; DB 1; Length 240;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 KTLNELK 135
Db 226 KTLNELK 232

RESULT 24
CLCR_RAT STANDARD; PRT; 268 AA.
ID CLCR_RAT
AC P55091; O63188;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Caldeactin precursor (EC 3.4.21.2) (Chymotrypsin C) (Serum calcium-
DE decreasing factor).
DE CTRC.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1]
RN
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Pancreas;
RX MEDLINE=96107178; PubMed=8530454;
RA Tomomura A., Tomomura M., Fukushima T., Akiyama M., Kubota N.,
RA Kumaki K., Nishii Y., Nohkura T., Saheki T.;
RT "Molecular cloning and expression of serum calcium-decreasing factor
RT (caldeactin).";
RL J. Biol. Chem. 270:30315-30321(1995).
[2]

```

RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=92165057; PubMed=1537555;
RA Kang U., Wiegand U., Mueller-Hill B.;
RT "Identification of cDNAs encoding two novel rat pancreatic serine
RL processes.";
RL Gene 110:181-187(1992).
RN (3)
RP CHARACTERIZATION.
RX MEDLINE=98207038; PubMed=9538241;
RA Yoshino-Yasuda I., Kobayashi K., Akiyama M., Tomomura A.,
RA Sabeji T.;
RT "Caldecrin is a novel-type serine protease expressed in pancreas, but
RT its homologue, elastase IV, is an artifact during cloning derived
RT from caldecrin gene.";
RL J. Biochem. 123:546-554(1998).
CC -1- FUNCTION: Has chymotrypsin-type protease activity and hypocalcemic
CC activity.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Leu|-Xaa, Tyr|-Xaa,
CC Phe|-Xaa, Met|-Xaa, Trp|-Xaa, Gln|-Xaa, Asn|-Xaa.
CC -1- TISSUE SPECIFICITY: Pancreas.
CC -1- SIMILARITY: Belongs to peptidase family S1. Elastase subfamily.
CC -1- CAUTION: Was originally (Ref.2) thought to be elastase IV.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S80379; AAB35830.1; -;
DR EMBL; X59014; CAA41753.1; -;
DR PIR; J01473; J01473.
DR HSP; P00766; ICHG.
DR MEROPS; S01.157; -;
DR InterPro; IPR009003; Cys_ser_trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; trypsin_1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; tryp-SFC; 1.
DR PROSITE; PS50240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Hydrolyse; Serine protease; Glycoprotein; Zymogen; Signal.
FT SIGNAL 1 16
FT PROPEP 17 29
FT CHAIN 30 268
FT ACT_SITE 74 74
FT ACT_SITE 123 123
FT ACT_SITE 216 216
FT DISULFID 17 141
FT DISULFID 59 75
FT DISULFID 155 222
FT DISULFID 186 202
FT DISULFID 212 243
FT CARBOHYD 25 25
FT CARBOHYD 90 90
FT CONFLICT 42 42
FT CONFLICT 96 120
SQ SEQUENCE 268 AA; 29374 MW; 33B67AF34D0F8583 CRC64;
TSMSGTDSSCGT (IN REF. 2).
TSMSGVAEDTVYHMKMLFLMN -> AEACTLRMTPS
Query Match 2.5%; Score 7; DB 1; Length 268;
Best Local Similarity 100.0%;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 25
ID_SAPF_ECOLI STANDARD; PRT; 268 AA.
AC P36637;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Peptide transport system ATP-binding protein sapf.
GN SAPF OR B1290 OR C1767 OR SF1295 OR S1377.
OS Escherichia coli,
OS Escherichia coli O6, and
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 217992, 623;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli;
RA Bergler H., Ebeling A., Fuchsbichler S., Hogenauer G., Turnowsky F.;
RL Submitted (Mar-1994) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12;
RA Epstein W., Noelker B., Stumpe S., Tewes R., Schmid R., Bakker E.P.;
RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12;
RX MEDLINE=9751357; PubMed=9097039;
RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M.,
RA Kitagawa M., Makino K., Nakade S., Nakamura Y., Nishimoto H., Nishio Y.,
RA Oshima T., Saito N., Sampei G., Seki Y., Sivasubram S.,
RA Tagami H., Takeda J., Takemoto K., Takeuchi Y., Wada C.,
RA Yamamoto Y., Horuchi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 28,040.1 min region on the linkage map.";
RL DNA Res. 3:363-377(1996).
RN [5]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=06:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Raeko D., Buckles E.L., Itoh S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
RN [6]
RP SEQUENCE FROM N.A.
RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441(2002).

```

RN [7]
RP SEQUENCE FROM N.A.
RX SPECIES=flexneri; STRAIN=2457T / ATCC 700930 / Serotype 2a;
R MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T."
RL Infect. Immun. 71:2775-2786(2003).
CC -1- FUNCTION: INVOLVED IN A PEPTIDE INTAKE TRANSPORT SYSTEM THAT
CC PLAYS A ROLE IN THE RESISTANCE TO ANTIMICROBIAL PEPTIDES.
CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
CC -1- SIMILARITY: Belongs to the ABC transporter family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U08190; AAA17671.1; -
DR EMBL; X97282; CA65941.1; -
DR EMBL; AE000227; AAC74372.1; -
DR EMBL; D90766; BAA14843.1; -
DR EMBL; D90767; BAA14851.1; -
DR EMBL; AE016760; AAN80233.1; -
DR EMBL; AE015156; AAN42906.1; -
DR EMBL; AE016982; AAP16789.1; -
DR PIR; E64877; E64877.
DR EcoGene; EG12305; sapF.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transporter.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 1.
KW Peptide transport; Transport; Inner membrane; ATP-binding;
KW Complete proteome.
FT NP_BIND 47
SQ SEQUENCE 268 AA; 30570 MW; F3B98BC9DC6CBFA CRC64;
FT
Query Match 2.5%; Score 7; DB 1; Length 268;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 149 LAPALIL 155
Db 161 LAPALIL 167

```

```

RT "Molecular genetic analysis of a locus required for resistance to
RT antimicrobial peptides in Salmonella typhimurium."
RL EMBL J. 12:4053-4062(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=LT2 / SGSC1412 / ATCC 700720;
R MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., All J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Kulkarny B.,
RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2."
RL Nature 413:852-856(2001).
CC -1- FUNCTION: INVOLVED IN A PEPTIDE INTAKE TRANSPORT SYSTEM THAT
CC PLAYS A ROLE IN THE RESISTANCE TO ANTIMICROBIAL PEPTIDES.
CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
CC -1- SIMILARITY: Belongs to the ABC transporter family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X74212; CA552268.1; -
DR EMBL; AE008775; AAL20613.1; -
DR PIR; G39589; S39589.
DR StyGene; SG10381; sapF.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transporter.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 1.
KW Peptide transport; Transport; Inner membrane; ATP-binding;
KW Complete proteome.
FT NP_BIND 47
SQ SEQUENCE 268 AA; 30671 MW; 339B649A75252504 CRC64;
FT
Query Match 2.5%; Score 7; DB 1; Length 268;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 149 LAPALIL 155
Db 161 LAPALIL 167

```

```

RESULT 26
SAPF_SALTY
ID SAPF_SALTY STANDARD; PRT; 268 AA.
AC P36638;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Peptide transport system ATP-binding protein sapF.
GN SAPF OR STM1696.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 14028S;
R MEDLINE=94038887; PubMed=8223423;

```

```

RESULT 27
SAPF_HABIN
ID SAPF_HABIN STANDARD; PRT; 269 AA.
AC P45289;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Peptide transport system ATP-binding protein sapF.
GN SAPF OR H11642.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN
RP SEQUENCE FROM N.A.
RX STRAIN=Rd / KW20 / ATCC 51907;
R MEDLINE=95350630; PubMed=7542800;
RA Fleisemann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Keriavage A.R., Bull C.J., Tomb J.-F., Dougherty B.A., Merrick J.W.,

```

RA Scott J.D., Shibley R., Liu L.-I., Glodok A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Utechtack T.R., Hanna M.C., Nguyen D.T., Saudak D.M., Brandon R.C.,
 RA Fine L.D., Fitchman J.L., Fuhrman J.L., Geoghagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.,
 RA "Whole-genome random sequencing and assembly of Haemophilus influenzae
 RT Rd.",
 RL Science 269:496-512(1995).
 CC -1- FUNCTION: INVOLVED IN A PEPTIDE INTAKE TRANSPORT SYSTEM THAT
 CC PLAYS A ROLE IN THE RESISTANCE TO ANTIMICROBIAL PEPTIDES
 CC (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
 CC -1- SIMILARITY: Belongs to the ABC transporter family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U32837; AAC23289.1; -.
 CC PIR: B64134; E64134.
 CC TIGR: H1642; -.
 CC InterPro: IPR003593; AAA_ATPase.
 CC InterPro: IPR003439; ABC_transporter.
 CC Pfam: PF00005; ABC_tran; 1.
 CC SMART: SM00362; AAA; 1.
 CC DR PROSITE: PS00211; ABC_TRANSPORTER_1; FALSE_NEG.
 CC DR PROSITE: PS00893; ABC_TRANSPORTER_2; 1.
 CC KW Peptide transport; Transport; Inner membrane; ATP-binding;
 CC Complete proteome.
 CC NP_BIND 45 ATP (POTENTIAL).
 CC SEQUENCE 269 AA; 30294 MW; E5C044EBCBDB801F CRC64;
 SQ
 Query Match 2.5%; Score 7; DB 1; Length 269;
 Best Local Similarity 100.0%; Pred. No. 44;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 149 LARLIL 155
 DB 159 LARLIL 165

RESULT 28
 ID METO_YERPE STANDARD; PRT; 271 AA.
 AC Q8ZHD0; 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE D-methionine-binding lipoprotein meto precursor.
 GN METO OR YPO1071 OR Y3106.
 OS Yersinia pestis.
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC Enterobacteriaceae; Yersinia.
 CC NCBI_TaxID=632;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN=CO-92 / Biovar Orientalis;
 CC MEDLINE=21470413; PubMed=11586360;
 CC PAKHILL U., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
 CC Prentice M.B., Sebaita M., James K.D., Churcher C., Mungall K.L.,
 CC Baker S., Baaham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,
 CC Chillingworth T., Cronin N., Davies R.M., Davis P., Dougan G.,
 CC Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
 CC Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.,
 CC Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.,
 CC "Genome sequence of Yersinia pestis, the causative agent of plague.",
 RL Nature 413:523-527(2001).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=K105 / Biovar Mediaevalis;
 RX MEDLINE=22137863; PubMed=12142430;
 RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
 RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
 RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
 RA Straley S.C., McDonough K.A., Nilles M.L., Watson J.S., Blattner F.R.,
 RA Perry R.D.,
 RL "Genome sequence of Yersinia pestis K105",
 RL J. Bacteriol. 184:4601-4611(2002).
 CC -1- FUNCTION: This protein is a component of a D-methionine
 CC permease, a binding protein-dependent, ATP-driven transport
 CC system (By similarity).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 CC (Probable).
 CC -1- MICELLANEOUS: The metIQ system is also to be able to transport
 CC the toxic methionine analog alpha-methyl-methionine (By
 CC similarity).
 CC -1- SIMILARITY: Belongs to the nlpA lipoprotein family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: A141416; CAC89914.1; -.
 CC EMBL: AE013911; AA86656.1; -.
 CC PIR: AG0131; AG0131.
 CC InterPro: IPR004872; Lipoprotein_9.
 CC InterPro: IPR000437; Prok_Lipoprot_S.
 CC InterPro: IPR004478; YaeC.
 CC DR Pfam: PF03180; Lipoprotein_9; 1.
 CC DR TIGRfam: TIGR00363; TIGR00363; 1.
 CC DR PROSITE: PS00013; PROKAR_LIPOPROTEIN_1.
 CC KW Transport; Amino-acid transport; Membrane; Lipoprotein; Signal;
 CC Complete proteome; Palmitate.
 CC FT SIGNAL 1 22 POTENTIAL.
 CC FT CHAIN 23 271 D-METHIONINE-BINDING LIPOPROTEIN METO.
 CC FT LIPID 23 23 N-palmitoyl cysteine (Potential).
 CC FT LIPID 23 23 S-diacylglycerol cysteine (Potential).
 CC SEQUENCE 271 AA; 29376 MW; 5214CEB0511F8C8 CRC64;
 SQ
 Query Match 2.5%; Score 7; DB 1; Length 271;
 Best Local Similarity 100.0%; Pred. No. 44;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 160 GLIKKD 166
 DB 153 GLIKKD 159

RESULT 29
 ID PLPB_PASHA STANDARD; PRT; 276 AA.
 AC Q08669; Q07364; 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Outer membrane lipoprotein 2 precursor (PLP2).
 GN PLPB.
 OS Pasteurella haemolytica.
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 CC Pasteurellaceae; Mannheimia.
 CC NCBI_TaxID=75985;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN=serotype A1;
 CC MEDLINE=9332810; PubMed=8335249;
 CC Murphy G.L., Whitworth L.C.;
 RA "Analysis of tandem, multiple genes encoding 30-kDa membrane proteins

RT in Pasteurella haemolytica A1.";
RL Gene 129:107-111(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Serotype A1;
RX MEDLINE=9401378; PubMed=8406866;
RA Cooney B.J., Lo R.Y.C.;
RT "Three contiguous lipoprotein genes in Pasteurella haemolytica A1
RT which are homologous to a lipoprotein gene in Haemophilus influenzae
RT type b.";
RL Infect. Immun. 61:4682-4688(1993).
CC -1- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
CC anchor (Probable).
CC -1- SIMILARITY: Belongs to the alpha lipoprotein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L11037; AAA2553.1; -;
DR EMBL; L16627; AAA2554.1; -;
DR PIR; JN0752; JN0752.
DR InterPro; IPR004872; Lipoprotein 9.
DR InterPro; IPR004437; Prok_lipoprot_S.
DR InterPro; IPR004478; YaeC
DR Pfam; PF03180; Lipoprotein 9; 1.
DR TIGRfam; TIGR00363; TIGR00363; 1.
DR PROSITE; PS00013; PROCAR_LIPOPROTEIN; 1.
KW Outer membrane; Lipoprotein; Signal; Multigene family; Palmitate.
FT SIGNAL 1 19
FT CHAIN 20 276
FT LIPID 20 20 N-palmitoyl cysteine (Probable).
FT LIPID 20 20 S-diacylglycerol cysteine (Probable).
FT CONFLICT 268 270 FNG -> LIVH (IN REF. 2).
SQ SEQUENCE 276 AA; 30157 MW; DF8F434ADD4D950 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 276;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 ELKDGAT 139
DB 133 ELKDGAT 139

RESULT 30
NAK2_ECOL6
ID NAK2_ECOL6 STANDARD; PRT; 291 AA.
AC Q8FDU8;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Putative N-acetylmannosamine kinase 2 (EC 2.7.1.60) (MannAC kinase 2).
GN NANK2 OR C9638.
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
[1]
RP SEQUENCE FROM N.A.
RX STRAIN=O6:HI / CFT073 / ATCC 700928;
RX MEDLINE=2388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G., III, Redford P., Roesch P.,
RA Maslow D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";

CC -1- FUNCTION: Catalyzes the phosphorylation of the N-acetylmannosamine
CC (MannAc) liberated from N-acetyl-neuraminic acid by the nank
CC protein (Potential).
CC -1- CATALYTIC ACTIVITY: ATP + N-acetyl-D-mannosamine = ADP + N-acetyl-D-
CC mannosamine 6-phosphate.
CC -1- PATHWAY: N-acetylmannosamine utilization.
CC -1- PATHWAY: N-acetylneuraminic acid (sialic acid) utilization.
CC -1- SIMILARITY: Belongs to the KOK (sugar kinases) family. Nank
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE016766; AAN82086.1; -;
DR HAMAP; MF_01234; -; 1.
DR InterPro; IPR006000; ROK.
DR Pfam; PF00480; ROK; 1.
DR PROSITE; PS01125; ROK; 1.
KW Carbohydrate metabolism; Transferase; Kinase; ATP-binding;
KW Complete proteome.
FT NP BIND 5 12
FT NP BIND 132 139
SQ SEQUENCE 291 AA; 30743 MW; 260BAD700BB34E4 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 291;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 LAGSTK 127
DB 191 LAGSTK 197

RESULT 31
ADT1_BOVIN
ID ADT1_BOVIN STANDARD; PRT; 297 AA.
AC P02722;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE ADP/ATP carrier protein, heart isoform T1 (ADP/ATP translocase 1)
DE (Adenine nucleotide translocator 1) (ANT 1).
GN SLC25A4 OR ANMT1
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89228093; PubMed=2540808;
RA Powell S.J., Medd S.M., Runswick M.J., Walker J.E.;
RA "Two bovine genes for mitochondrial ADP/ATP translocase expressed
RT differences in various tissues.";
RL Biochemistry 28:866-873(1989).
[2]
RP SEQUENCE
RX MEDLINE=82188267; PubMed=7076130;
RA Aquila H., Mista D., Eulitz M., Klingenberg M.;
RT "Complete amino acid sequence of the ADP/ATP carrier from beef heart
RT mitochondria.";
RL Hoppe-Seyler's Z. Physiol. Chem. 363:345-349(1982).
[3]
RP SEQUENCE OF 207-297 FROM N.A.
RX MEDLINE=86295775; PubMed=3017341;
RA Rasmussen U.B., Wohlrad H.;
RT "Bovine cardiac mitochondrial ADP/ATP-carrier: two distinct mRNAs and

RL Biochem. Biophys. Res. Commun. 138:850-857(1986).

CC -1- FUNCTION: Catalyzes the exchange of ADP and ATP across the mitochondrial inner membrane.

CC -1- SUBUNIT: Homodimer.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial inner membrane.

CC -1- SIMILARITY: Contains 3 Solcar repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC

CC EMBL: M13783; AAA30363.1; -

CC EMBL: M24102; AAA30766.1; -

CC PIR: A43646; XMO.

CC InterPro: IPR002067; Mit_carrier.

CC InterPro: IPR002030; Mit_uncoupling.

CC InterPro: IPR001993; Mitoch_carrier.

CC Pfam: PF00153; mito_carr; 3.

CC PRINTS: PR00926; MITOCARRIER.

CC PRINTS: PR00784; MTUNCOUPLING.

CC PROSITE: PS50920; SOLCAR; 3.

CC Mitocondrion; Inner membrane; Repeat; Transmembrane; Transport; K+ MultiGene family; Methylation.

CC INIT MET 0

CC TRANSSEM 11 28 1 (POTENTIAL).

CC TRANSSEM 72 90 2 (POTENTIAL).

CC TRANSSEM 116 133 3 (POTENTIAL).

CC TRANSSEM 175 194 4 (POTENTIAL).

CC TRANSSEM 213 230 5 (POTENTIAL).

CC TRANSSEM 272 290 6 (POTENTIAL).

CC REPEAT 5 97 SOLCAR 1.

CC REPEAT 110 200 SOLCAR 2.

CC REPEAT 211 296 SOLCAR 3.

CC MOD_RES 1 1 BLOCKED.

CC FT MOD_RES 51 51 METHYLATION (POTENTIAL).

CC SQ SEQUENCE 297 AA; 32836 MW; A582D3C4A0A8B48 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 297;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 29 AAISKTA 35
| | | | |
DB 18 AAISKTA 24

RESULT 32

YD67_METTA STANDARD; PRT; 297 AA.

AC Q58762; 15-JUL-1999 (Rel. 38, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Hypothetical ABC transporter ATP-binding protein MJ1367.

GN MJ1367.

OS Methanococcus jannaschii.

OC Archaea; Euryarchaeota; Methanococci; Methanococcales;

OC Methanocaldococcaceae; Methanocaldococcus.

OX NCBI_TaxID=2190;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;

RX MEDLINE=96337999; PubMed=66889877; Zhou L., Fleischmann R.D., Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D., Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D., Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I., Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodex A.,

RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D., Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C., Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M., Kleck H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C., "Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii." Science 273:1058-1073(1996).

CC -1- FUNCTION: Might be part of an ABC transporter complex. Might be responsible for energy coupling to the transport system.

CC -1- SUBUNIT: Might form a complex with the permease protein MJ1366.

CC -1- SIMILARITY: Belongs to the ABC transporter family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC

CC EMBL: U67576; AAB99375.1; -

CC PIR: F64470; F64470.

CC HSPD: Q58663; 16GH.

CC TIGR: MJ1367; -

CC InterPro: IPR003593; AAA_ATPase.

CC InterPro: IPR004339; ABC_transporter.

CC Pfam: PF00005; ABC_tran; 1.

CC Prodom: PD000006; ABC_transporter; 1.

CC SMART: SM00382; AAA; 1.

CC PROSITE: PS00211; ABC_TRANSPORTER_1; 1.

CC PROSITE: PS50893; ABC_TRANSPORTER_2; 1.

CC Hypothetical protein; ATP-binding; Transport; Complete proteome.

CC NP_BIND 32 39 ATP (POTENTIAL).

CC SQ SEQUENCE 297 AA; 33266 MW; 710DB22733EBB12 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 297;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 149 LARALIL 155
| | | | |
DB 140 LARALIL 146

RESULT 33

ADT2_HUMAN STANDARD; PRT; 298 AA.

AC P05141; Q43350; 13-AUG-1987 (Rel. 05, Created)

DT 01-OCT-1994 (Rel. 30, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE ADP/ATP carrier protein, fibroblast isoform (ADP/ATP translocase 2)

DE Adenine nucleotide translocator 2 (ANT 2).

GN SLC25A5 OR ANT2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Placenta;

RX MEDLINE=90375457; PubMed=2168878;

RA Ku D.-H., Kagan J., Chen S.-T., Chang C.-D., Baserga R., Wurzel J., "The human fibroblast adenine nucleotide translocator gene. Molecular cloning and sequence." J. Biol. Chem. 265:16060-16063(1990).

RL [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=87166056; PubMed=3031073;

RA Battini R., Ferreri S., Kaczmarek L., Calabretta B., Chen S.T., Baserga R., "Molecular cloning of a cDNA for a human ADP/ATP carrier which is growth-regulated."

RN J. Biol. Chem. 262:4355-4358 (1987).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Chen C.N., Su Y., Baybayan P., Siruno A., Nagaraja R.,
 RA Mazzarella R.A., Schlesinger D., Chen E.Y.,
 RN Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Becker M., Graves T., Ozersky P.,
 RN Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Eye;
 RX MEDLINE=23388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schlier G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang U., Hsieh F.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares W.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.T., Abramson R.D., Miliuk S.J.,
 RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gamarathne P.H.,
 RA Richards S., Wooley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Rahy J., Helton E., Kettelman K., Madan A., Rodriguez S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Scherach A., Schein J.E., Jones S.J.M., Marra M.A.,
 RA "generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences."
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [6]
 RN SEQUENCE OF 47-298 FROM N.A.
 RP TISSUE=Liver;
 RX MEDLINE=88124845; PubMed=2829183;
 RA Houldsworth J., Attardi G.,
 RT "Two distinct genes for ADP/ATP translocase are expressed at the mRNA
 level in adult human liver."
 RT Proc. Natl. Acad. Sci. U.S.A. 85:377-381 (1988).
 CC -1- FUNCTION: Catalyzes the exchange of ADP and ATP across the
 CC mitochondrial inner membrane.
 CC -1- SUBUNIT: Homodimer.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC inner membrane.
 CC -1- SIMILARITY: Belongs to the mitochondrial carrier family.
 CC -1- SIMILARITY: Contains 3 Solcar repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M57434; AAA51737.1; -
 DR EMBL: J02683; AAA5579.1; -
 DR EMBL: U78810; AAB39266.1; -
 DR EMBL: AC004000; AAB96347.1; -
 DR EMBL: BC056160; AAB56160.1; -
 DR EMBL: J03591; AAA36749.1; -
 DR PIR: A29132; A29132.
 DR GeneW: HGNC:10991; SLC25A5.
 DR MIM: 300150.
 DR GO: GO:0008877; C: integral to plasma membrane; TAS.
 DR GO: GO:0015207; F: adenine transporter activity; TAS.
 DR GO: GO:0006810; P: transport; TAS.
 DR InterPro: IPR002067; M: carrier.
 DR InterPro: IPR002030; M: uncoupling.
 DR InterPro: IPR001993; M: mitochondrial carrier.
 DR Gene: 254:57-66 (2000).

DR Pfam: PF00153; mito_carr. 3.
 DR PRINTS; PRO0926; MITOCARRIER.
 DR PRINTS; PRO0784; MITOUNCOUPLING.
 DR PROSITE; PS50920; SOLCAR; 3.
 KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport;
 KW Multigene family.
 FT TRANSMEM 12 29 1 (POTENTIAL).
 FT TRANSMEM 73 91 2 (POTENTIAL).
 FT TRANSMEM 117 134 3 (POTENTIAL).
 FT TRANSMEM 176 195 4 (POTENTIAL).
 FT TRANSMEM 214 231 5 (POTENTIAL).
 FT TRANSMEM 273 291 6 (POTENTIAL).
 FT REPEAT 6 98 SOLCAR 1.
 FT REPEAT 111 201 SOLCAR 2.
 FT REPEAT 212 297 SOLCAR 3.
 FT CONFLICT 6 6 V -> L (IN REF. 2).
 FT CONFLICT 66 66 G -> E (IN REF. 2).
 FT CONFLICT 111 111 R -> L (IN REF. 4 AND 6).
 FT CONFLICT 162 162 V -> G (IN REF. 6).
 SQ SEQUENCE 298 AA; 32895 MW; F973C3AED92C49D3 CRC64;
 Query Match 2.5% Score 7; DB 1; Length 298;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 29 AAIKTKA 35
 DB 19 AAIKTKA 25
 RESULT 34
 ADIT_MOUSE STANDARD; PRT; 298 AA.
 AC P51881; O61311;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE ADP/ATP carrier protein, fibroblast isoform (ADP/ATP translocase 2)
 DE (Adenine nucleotide translocator 2) (ANT 2).
 GN SLC25A5 OR ANT2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=97059403; PubMed=8903724;
 RA Ellison J.W., Li X., Franke U., Shapiro L.J.,
 RT "Rapid evolution of human pseudautosomal genes and their mouse
 RT homologs."
 RT Mamm. Genome 7:25-30 (1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skeletal muscle;
 RA Sheldon J.G.;
 RN Thesis (1995), University of Cambridge, U.K.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/Sv;
 RA Coester P., Laplace C.;
 RN Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP REVISIONS.
 RA Laplace C.;
 RN Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20432087; PubMed=10974536;
 RA Levy S.E., Chen Y.-S., Graham B.H., Wallace D.C.;
 RT "Expression and sequence analysis of the mouse adenine nucleotide
 RT translocase 1 and 2 genes."
 RN Gene 254:57-66 (2000).

RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Mammary gland;
 RX MEDLINE=2238825; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marisana K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Pandey C.,
 RA Raha S.S., Loquellano N.A., Peters G.U., Abramson R.D., Mallaby S.J.,
 RA Bosak S.A., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fehey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc Natl Acad Sci U S A. 99:16899-16903(2002).
 CC -!- FUNCTION: Catalyzes the exchange of ADP and ATP across the
 CC mitochondrial inner membrane.
 CC -!- SUBUNIT: Homodimer.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC inner membrane.
 CC -!- SIMILARITY: Belongs to the mitochondrial carrier family.
 CC -!- SIMILARITY: Contains 3 Solcar repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U27316; AAC52838.1; -;
 DR EMBL: U10404; AAA19009.1; -;
 DR EMBL: X70847; CA50196.1; -;
 DR EMBL: AF240003; AAF64471.1; -;
 DR EMBL: BC004570; AA04570.1; -;
 DR WGI: MGI:1353496; SLOC25A5.
 DR InterPro: IPR002067; Mlt_carrier.
 DR InterPro: IPR002030; Mlt_uncoupling.
 DR InterPro: IPR001993; Mitoch_carrier.
 DR Pfam: PF00153; mltcarr; 3.
 DR PRINTS: PR00926; MITOCARRIER.
 DR PROSITE: PS00784; MITOCOUPLING.
 DR PRINTS: PR00784; MITOCOUPLING.
 DR PROSITE: PS00920; SOLCAR; 3.
 KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport;
 KW Multigene family.
 KM TRANSMEM 12 29 1 (POTENTIAL).
 FT TRANSMEM 73 91 2 (POTENTIAL).
 FT TRANSMEM 117 134 3 (POTENTIAL).
 FT TRANSMEM 176 195 4 (POTENTIAL).
 FT TRANSMEM 214 231 5 (POTENTIAL).
 FT TRANSMEM 273 291 6 (POTENTIAL).
 FT REPEAT 6 98 SOLCAR 1.
 FT REPEAT 111 201 SOLCAR 2.
 FT REPEAT 212 297 SOLCAR 3.
 SQ SEQUENCE 298 AA; 32931 MM; 0798804B987EFE20 CRC64;
 QY Query Match 2.5%; Score 7; DB 1; Length 298;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 29 AAISKTA 35
 19 AAISKTA 25

RESULT 35
 ID ADT2 RAT STANDARD; PRT; 298 AA.
 AC 009073;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE ADP/ATP carrier protein, fibroblast isoform (ADP/ATP translocase 2)
 DE (Adenine nucleotide translocator 2) (ANT 2).
 CN SLOC25A5 OR ANT2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_Taxid=10116;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Liver;
 RX MEDLINE=94002161; PubMed=8399300;
 RA Shiohara Y., Kamida M., Yamazaki N., Terada H.,
 RA "Isolation and characterization of cDNA clones and a genomic clone
 RT encoding rat mitochondrial adenine nucleotide translocator.";
 RL Biochim. Biophys. Acta 1152:192-196(1993).
 CC -!- FUNCTION: Catalyzes the exchange of ADP and ATP across the
 CC mitochondrial inner membrane.
 CC -!- SUBUNIT: Homodimer.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC inner membrane.
 CC -!- TISSUE SPECIFICITY: Present in kidney, brain, heart, liver and
 CC skeletal muscle.
 CC -!- SIMILARITY: Belongs to the mitochondrial carrier family.
 CC -!- SIMILARITY: Contains 3 Solcar repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D12771; BA00238.1; -;
 DR InterPro: IPR002067; Mlt_carrier.
 DR InterPro: IPR002030; Mlt_uncoupling.
 DR InterPro: IPR001993; Mitoch_carrier.
 DR Pfam: PF00153; mltcarr; 3.
 DR PRINTS: PR00926; MITOCARRIER.
 DR PROSITE: PS00784; MITOCOUPLING.
 DR PROSITE: PS00920; SOLCAR; 3.
 KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport;
 KW Multigene family.
 KM TRANSMEM 12 29 1 (POTENTIAL).
 FT TRANSMEM 73 91 2 (POTENTIAL).
 FT TRANSMEM 117 134 3 (POTENTIAL).
 FT TRANSMEM 176 195 4 (POTENTIAL).
 FT TRANSMEM 214 231 5 (POTENTIAL).
 FT TRANSMEM 273 291 6 (POTENTIAL).
 FT REPEAT 6 98 SOLCAR 1.
 FT REPEAT 111 201 SOLCAR 2.
 FT REPEAT 212 297 SOLCAR 3.
 SQ SEQUENCE 298 AA; 32901 MM; 6A59204B987EFE35 CRC64;
 QY Query Match 2.5%; Score 7; DB 1; Length 298;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 29 AAISKTA 35
 19 AAISKTA 25

RESULT 36

ADT3_BOVIN STANDARD; PRT; 298 AA.
ID ADT3_BOVIN
AC P32007;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE ADP,ATP carrier protein, isoform T2 (ADP/ATP translocase 3) (adenine nucleotide translocator 3) (ANT 3).
GN SLC25A6 OR ANT3.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89229093; PubMed=2540808;
RA Powell S.J., Mead S.M., Runswick M.J., Walker J.E.;
RT "Two bovine genes for mitochondrial ADP/ATP translocase expressed differences in various tissues.";
RL Biochemistry 28:866-873(1989).
CC -1- FUNCTION: Catalyzes the exchange of ADP and ATP across the mitochondrial inner membrane.
CC -1- SUBUNIT: Homodimer.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial inner membrane.
CC -1- SIMILARITY: Belongs to the mitochondrial carrier family.
CC -1- SIMILARITY: Contains 3 Solcar repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M24103; AAA30769.1; -
DR PIR; B43646; B43646.
DR InterPro; IPR002067; Mit_carrier.
DR InterPro; IPR002030; Mit_uncoupling.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PFO0153; mito_carri_3.
DR PRINTS; PR00926; MITOCARRIER.
DR PRINTS; PR00784; MITOCOUPLING.
DR PROSITE; PS50920; SOLCAR_3.
KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport; MultiGene Family.
KM
FT TRANSMEM 12 29 1 (POTENTIAL).
FT TRANSMEM 73 91 2 (POTENTIAL).
FT TRANSMEM 117 134 3 (POTENTIAL).
FT TRANSMEM 176 195 4 (POTENTIAL).
FT TRANSMEM 214 231 5 (POTENTIAL).
FT TRANSMEM 273 291 6 (POTENTIAL).
FT REPEAT 6 98 SOLCAR 1.
FT REPEAT 111 201 SOLCAR 2.
FT REPEAT 212 297 SOLCAR 3.
SQ SEQUENCE 298 AA; 32877 MW; 1C34E7DF6E4061 CRC64;
Query Match 2.5%; Score 7; DR 1; Length 298;
Best Local Similarity 100.0%; Fred No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ADT3_HUMAN STANDARD; PRT; 298 AA.
ID ADT3_HUMAN
AC P1236; O96C49;
DT 01-OCT-1989 (Rel. 12, Created)
01-NOV-1990 (Rel. 16, Last sequence update)
10-OCT-2003 (Rel. 42, Last annotation update)
DE ADP,ATP carrier protein, liver isoform T2 (ADP/ATP translocase 3) (adenine nucleotide translocator 3) (ANT 3).
GN SLC25A6 OR ANT3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89236396; PubMed=2541251;
RA Cozens A.L., Runswick M.J., Walker J.E.;
RT "DNA sequences of two expressed nuclear genes for human mitochondrial ADP/ATP translocase.";
RL J. Mol. Biol. 206:261-280(1989).
CC [2]
RP SEQUENCE FROM N.A.
RX Zhou U., Yu W., Tang H., Mei G., Tsang Y.T.M., Bouck J., Gibbs R.A., Margolin J.F.;
RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC [3]
RP TISSUE=Brain, Cervix, Eye, and Lung;
RX MEDLINE=22388257; PubMed=12477932;
RA Klausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Altschul S.F., Zeeberg B., Buelow K.H., Schaefer C.F., Bat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Mang J., Hsieh F., Diatchenko L., Marasini K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Cassavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carminci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulihy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunnarsson P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko I., Boulfard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences".
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP SEQUENCE OF 36-298 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=88124845; PubMed=2829183;
RA Houldsworth J., Attardi G.;
RT "Two distinct genes for ADP/ATP translocase are expressed at the mRNA level in adult human liver.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:377-381(1988).
CC -1- FUNCTION: Catalyzes the exchange of ADP and ATP across the mitochondrial inner membrane.
CC -1- SUBUNIT: Homodimer.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial inner membrane.
CC -1- SIMILARITY: Belongs to the mitochondrial carrier family.
CC -1- SIMILARITY: Contains 3 Solcar repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J03592; AAA36750.1; -
DR EMBL; AY007135; AA01998.1; -
DR EMBL; BC007295; AA07295.1; -
DR EMBL; BC007850; AA07850.1; -

DR EMBL: BC008737; AAH08737.1; -
 DR EMBL: BC008935; AAH08935.1; -
 DR EMBL: BC014775; AAH14775.1; -
 DR EMBL: BC01912; AAH1912.1; -
 DR PIR: S03894; S03894.
 DR Genem: HGNC:10992; SLC25A6.
 DR MIM: 300151; -
 DR MIM: 403000; -
 DR CO: GO:0005744; Mitochondrial inner membrane pre-sequence t. . .; TMS.
 DR GO: GO:0005741; P-ATP/ADP antiporter activity; NAS.
 DR GO: GO:0006854; P-ATP/ADP exchange; TMS.
 DR InterPro: IPR002067; Mlt_carrier.
 DR InterPro: IPR002030; Mltc_carrier.
 DR Pfam: PF00153; mltc_carr; 3-
 DR Pfam: PF00926; MITOCOUPLING.
 DR PRINTS: PR00784; MITOCOUPLING.
 DR PROSITE: PS00920; SOLCAR; 3.
 DR KEGG: K01111; Mitochondrial inner membrane; Repeat; Transmembrane; Transport;
 KM Multigene family.
 FT TRANSMEM 12 29 1 (POTENTIAL).
 FT TRANSMEM 73 91 2 (POTENTIAL).
 FT TRANSMEM 117 134 3 (POTENTIAL).
 FT TRANSMEM 176 195 4 (POTENTIAL).
 FT TRANSMEM 214 231 5 (POTENTIAL).
 FT TRANSMEM 273 291 6 (POTENTIAL).
 FT REPEAT 6 98 SOLCAR 1.
 FT REPEAT 111 201 SOLCAR 2.
 FT REPEAT 212 297 SOLCAR 3.
 FT REPEAT 105 108 SOLCAR 3.
 FT CONFLICT 242 242 KHTO -> RHA (IN REF. 4).
 FT CONFLICT 242 242 S -> F (IN REF. 3); AAH14775).
 SQ SEQUENCE 298 AA; 32866 MW; 1853459F0B9672F CRC64;

Query Match 2.5%; Score 7; DB 1; Length 298;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 AAIKSTA 35
 DB 19 AAIKSTA 25

RESULT 38
 HEM3_NEIMA STANDARD; PRT; 311 AA.
 ID HEM3_NEIMA
 AC Q9UVS4;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Porphobilinogen deaminase (EC 2.5.1.61) (PBG) (Hydroxymethylbilane
 synthase) (HMB) (Pre-urophorphyrinogen synthase).
 GN HEMC OR NMA0718.
 OS Neisseria meningitidis (serogroup A).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC Neisseriaceae; Neisseria.
 CC NCB1_TaxID=65699;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=22491 / Serogroup A / Serotype 4A;
 RX MEDLINE=20222555; PubMed=10761919;
 RA Pakshill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
 Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
 Davies R.M., Davis P., Devlin K., Feltham T., Hamlin N., Holroyd S.,
 Jasek K., Leather S., Moule S., Mungall K., Quail M.A.,
 Rajandream M.A., Rutherford K.M., Simmonds W., Skelton J.,
 Whitehead S., Spratt B.G., Barrall B.G.;
 RT "Complete DNA sequence of a serogroup A strain of Neisseria
 meningitidis 22491.";
 RL Nucleic Acids Res. 29:22491-22499 (2001).
 CC -1- FUNCTION: Tetrapolymerization of the monopyrrole PBG into the
 hydroxymethylbilane preuroporphyrinogen in several discrete steps.
 CC -1- CATALYTIC ACTIVITY: 4 porphobilinogen + H(2)O =
 -1- CATALYTIC ACTIVITY: 4 porphobilinogen + H(2)O =
 hydroxymethylbilane + 4 NH(3).
 CC -1- CATALYTIC ACTIVITY: 4 porphobilinogen + H(2)O =
 hydroxymethylbilane + 4 NH(3).

CC -1- COFACTOR: Covalently binds a dipyrromethane cofactor to which the
 CC porphobilinogen subunits are added (By similarity).
 CC -1- PATHWAY: Porphyrin biosynthesis by the C5 pathway; Fourth step.
 CC -1- SUBUNIT: Monomer (By similarity).
 CC -1- SIMILARITY: Belongs to the HMB family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL, institution
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AL162754; CAB84003.1; -
 CC PIR: A81915; A81915.
 CC HSSP: P06983; LPDA.
 CC HAMAP: MF_00260; -; 1.
 CC InterPro: IPR000860; Porphobil_deam.
 CC Pfam: PF01379; Porphobil_deam.
 CC Pfam: PF03900; Porphobil_deam; 1.
 CC PRINTS: PR00151; PORPHOBILINASE.
 CC PRODOM: PD02745; Porphobil_deam; 1.
 CC TIGRFAMs: TIGR00212; hemc; 1.
 CC DR PROSITE: PS00533; PORPHOBILINASE DEAM; 1.
 KW Porphyrin biosynthesis; Transferase; Complete proteome.
 FT BINDING 242 242 PYRROMETHANE COFACTOR (BY SIMILARITY).
 SQ SEQUENCE 311 AA; 33448 MW; DFC7B5C6BFD3A03 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 311;
 Best Local Similarity 100.0%; Pred. No. 50;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 NPKXVI 186
 DB 2 NPKXVI 8

RESULT 39
 HEM3_NEIMA STANDARD; PRT; 311 AA.
 ID HEM3_NEIMA
 AC Q9KOF6;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Porphobilinogen deaminase (EC 2.5.1.61) (PBG) (Hydroxymethylbilane
 synthase) (HMB) (Pre-urophorphyrinogen synthase).
 GN HEMC OR NMA0539.
 OS Neisseria meningitidis (serogroup B).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC Neisseriaceae; Neisseria.
 CC NCB1_TaxID=491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MC58 / Serogroup B;
 RX MEDLINE=20175755; PubMed=10710307;
 RA Tetzelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
 Eissen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
 Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Dougherty B.A.,
 Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
 Mason T., Ciecko A., Parksey D.S., Blair E., Clifton H., Clark E.B.,
 Cotton M.D., Uitterback T.R., Khouri H., Qin H., Vamathevan J.,
 Gill J., Scarlato V., Nisiganti V., Pizze M., Grandi G., Sun L.,
 Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.,
 RA "Complete genome sequence of Neisseria meningitidis serogroup B strain
 MC58.";
 RT Science 287:1809-1815 (2000).
 CC -1- FUNCTION: Tetrapolymerization of the monopyrrole PBG into the
 hydroxymethylbilane preuroporphyrinogen in several discrete steps.
 CC -1- CATALYTIC ACTIVITY: 4 porphobilinogen + H(2)O =
 hydroxymethylbilane + 4 NH(3).
 CC -1- COFACTOR: Covalently binds a dipyrromethane cofactor to which the
 porphobilinogen subunits are added (By similarity).

```
CC -1- PATHWAY: Porphyrin biosynthesis by the C5 pathway; fourth step.
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SIMILARITY: Belongs to the HMBs family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE002409; AAF40968.1; -.
CC PIR; B81188; B81188.
CC HSSP; P06983; 1PDA.
CC TIGR; NMB0533; -.
CC HAVAP; MF_00260; -; 1.
CC InterPro; IPR000860; Porphobil deam.
CC Pfam; PF01379; Porphobil deam; 1.
CC Pfam; PF03900; Porphobil deam; 1.
CC PRINTS; PR00151; PORPHBDMASE.
CC PRODOM; PD002745; Porphobil deam; 1.
CC TIGRFAMs; TIGR00212; HemeC; 1.
CC PROSITE; PS00533; PORPROBILINGEN_DEAM; 1.
CC Porphyrin biosynthesis; Transferase; Complete proteome.
CC BINDING; 242 242 PYRROMETHANE COPACTOR (BY SIMILARITY).
CC SEQUENCE 311 AA; 33478 MW; 8446BBD7BF0577F CRC64;
SQ
Query Match 2.5%; Score 7; DB 1; Length 311;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 180 NPKKLV 186
Db 2 NPKKLV 8
RESULT 40
TFBI_THEAC STANDARD; PRT; 312 AA.
ID TFBI_THEAC
AC Q9HJN7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Transcription initiation factor IIB 1 (TFIIB 1).
GN TFBA OR TA0940.
OS Thermoplasma acidophilum.
OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
OC Thermoplasmataceae; Thermoplasma.
OX NCB1_TaxID=2303;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=DSK 1728;
RX MEDLINE=20479972; PubMed=11029001;
RA Ruepp A., Grael W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
RA Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
RA "The genome sequence of the thermoacidophilic scavenger Thermoplasma
RA acidophilum."
RT Nature 407:508-513(2000).
RL
CC -1- FUNCTION: Stabilizes TBP binding to an archaeal box-A promoter.
CC Also responsible for recruiting RNA polymerase II to the pre-
CC initiation complex (DNA-TBP-TFIIB) (By similarity).
CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -1- SIMILARITY: Belongs to the TFIIB family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

```
DR EMBL; AL445066; CAC12069.1; -.
DR HSSP; P29095; 1A1S.
DR HAVAP; MF_00383; -; 1.
DR InterPro; IPR006670; Cyclin.
DR InterPro; IPR000812; TFIIB euk.
DR Pfam; PF00382; transactin fac2; 2.
DR PRINTS; PR00685; TIFACTO1B.
DR SMART; SM00385; CYCLIN; 2.
DR PROSITE; PS00782; TFIIB; 2.
DR Transcription regulation; Repeat; Zinc-finger; Metal-binding; Zinc;
KW Complete proteome.
FT ZN FING 16 38 ZN-RIBBON TFIIB-TYPE.
FT REPEAT 129 212 1.
FT METAL 223 304 2.
FT METAL 16 16 ZINC (BY SIMILARITY).
FT METAL 19 16 ZINC (BY SIMILARITY).
FT METAL 35 35 ZINC (BY SIMILARITY).
FT METAL 38 38 ZINC (BY SIMILARITY).
SQ SEQUENCE 312 AA; 35100 MW; 85BE5459CA526837 CRC64;
Query Match 2.5%; Score 7; DB 1; Length 312;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 50 AVAEVAG 56
Db 283 AVAEVAG 289
```

Search completed: June 16, 2004, 11:16:05
Job time : 29 secs

1000

1000

1000

1000

1000

1000

1000

Mon Jun 22 11:04:02 2004

us-10-018-672-2.oligo.rapb

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 16, 2004, 11:17:18 ; Search time 48 Seconds
(without alignments)
1619.925 Million cell updates/sec

Title: US-10-018-672-2
Perfect score: 276
Sequence: 1 NMFNGKINGIALSGIALAG.....TDEVEAKKQFMDGVIKGM 276

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1158786 seqs, 281726120 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1158786

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

Published Applications: AA.*
1: /cgn2_6/ptodata/2/pubppa/US07_PUBCOMP.pep.*
2: /cgn2_6/ptodata/2/pubppa/PTCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubppa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubppa/US06_PUBCOMP.pep.*
5: /cgn2_6/ptodata/2/pubppa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubppa/PTCTUS_PUBCOMP.pep.*
7: /cgn2_6/ptodata/2/pubppa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubppa/US08_PUBCOMP.pep.*
9: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMP.pep.*
10: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMP.pep.*
11: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMP.pep.*
12: /cgn2_6/ptodata/2/pubppa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMP.pep.*
14: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMP.pep.*
15: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMP.pep.*
16: /cgn2_6/ptodata/2/pubppa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubppa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubppa/US60_PUBCOMP.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the chance being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	243	88.0	276	12 US-10-282-122A-6336	Sequence 63386, A
2	19	6.9	273	9 US-09-815-242-11067	Sequence 11067, A
3	19	6.9	273	12 US-10-282-122A-58240	Sequence 58240, A
4	19	6.9	273	12 US-10-380-817-2	Sequence 2, Appl1
5	19	6.9	273	12 US-10-380-817-4	Sequence 4, Appl1
6	19	6.9	273	12 US-10-380-817-6	Sequence 6, Appl1
7	19	6.9	273	12 US-10-380-817-8	Sequence 8, Appl1
8	19	6.9	273	12 US-10-380-817-10	Sequence 10, Appl1
9	16	5.8	271	12 US-10-282-122A-68391	Sequence 68391, A
10	13	4.7	270	12 US-10-282-122A-52838	Sequence 52838, A
11	12	4.3	276	12 US-10-282-122A-67387	Sequence 67387, A
12	11	4.0	67	13 US-10-092-243A-15587	Sequence 15, Appl1
13	11	4.0	271	12 US-10-282-122A-55926	Sequence 55926, A
14	11	4.0	271	12 US-10-282-122A-76195	Sequence 76195, A
15	10	3.6	276	12 US-10-282-122A-44792	Sequence 44792, A

16	9	3.3	241	12 US-10-282-122A-44788	Sequence 44788, A
17	9	3.3	256	12 US-10-282-122A-61127	Sequence 61127, A
18	9	3.3	262	12 US-10-282-122A-51371	Sequence 51371, A
19	9	3.3	270	12 US-10-282-122A-55804	Sequence 55804, A
20	9	3.3	273	12 US-10-282-122A-60558	Sequence 60558, A
21	9	3.3	275	12 US-10-282-122A-77174	Sequence 77174, A
22	8	2.9	96	9 US-09-764-853-805	Sequence 805, App
23	8	2.9	96	14 US-10-091-488-162	Sequence 162, App
24	8	2.9	143	12 US-10-424-559-262926	Sequence 262926, A
25	8	2.9	240	9 US-09-815-242-11656	Sequence 11656, A
26	8	2.9	263	12 US-10-282-122A-52976	Sequence 52976, A
27	8	2.9	265	12 US-10-282-122A-51201	Sequence 51201, A
28	8	2.9	270	12 US-10-282-122A-49339	Sequence 49339, A
29	8	2.9	271	9 US-09-815-242-10050	Sequence 10050, A
30	8	2.9	271	12 US-10-282-122A-43291	Sequence 43291, A
31	8	2.9	271	12 US-10-282-122A-59600	Sequence 59600, A
32	8	2.9	281	12 US-10-282-122A-74389	Sequence 74389, A
33	8	2.9	282	12 US-10-282-122A-57911	Sequence 57911, A
34	8	2.9	295	12 US-10-282-122A-47704	Sequence 47704, A
35	8	2.9	379	9 US-09-764-853-621	Sequence 621, App
36	8	2.9	467	15 US-10-369-493-20865	Sequence 20865, A
37	8	2.9	641	15 US-10-369-493-15592	Sequence 15592, A
38	8	2.9	739	9 US-09-854-845-45	Sequence 45, Appl
39	8	2.9	744	9 US-09-854-845-43	Sequence 43, Appl
40	8	2.9	766	9 US-09-854-845-49	Sequence 49, Appl
41	8	2.9	771	9 US-09-854-845-47	Sequence 47, Appl
42	8	2.9	838	9 US-09-854-845-23	Sequence 23, Appl
43	8	2.9	838	15 US-10-150-115-49	Sequence 49, Appl
44	8	2.9	838	15 US-10-369-072-49	Sequence 27, Appl
45	8	2.9	843	9 US-09-854-845-27	Sequence 27, Appl
46	8	2.9	865	9 US-09-854-845-33	Sequence 33, Appl
47	8	2.9	870	9 US-09-854-845-31	Sequence 31, Appl
48	7	2.5	69	12 US-10-424-599-195672	Sequence 195672, A
49	7	2.5	69	12 US-10-424-599-200956	Sequence 200956, A
50	7	2.5	73	12 US-10-335-977-4881	Sequence 4881, App
51	7	2.5	85	12 US-10-424-599-271432	Sequence 271432, A
52	7	2.5	87	12 US-10-424-599-271850	Sequence 271850, A
53	7	2.5	92	12 US-10-296-115-1035	Sequence 1035, App
54	7	2.5	94	12 US-10-424-599-212355	Sequence 212355, A
55	7	2.5	131	9 US-09-925-301-1808	Sequence 1808, App
56	7	2.5	132	9 US-09-925-301-1459	Sequence 1459, App
57	7	2.5	133	14 US-10-289-642-12	Sequence 12, Appl
58	7	2.5	134	12 US-10-424-599-163134	Sequence 163134, A
59	7	2.5	146	12 US-10-424-599-167109	Sequence 167109, A
60	6	2.5	153	12 US-10-335-977-9642	Sequence 9642, App
61	6	2.5	154	14 US-10-289-642-14	Sequence 14, Appl
62	6	2.5	156	12 US-10-424-599-205606	Sequence 205606, A
63	6	2.5	158	12 US-10-425-114-53426	Sequence 53426, A
64	6	2.5	159	9 US-09-963-791-18	Sequence 18, Appl
65	6	2.5	159	12 US-10-419-276-18	Sequence 18, Appl
66	6	2.5	174	14 US-10-017-161-700	Sequence 700, App
67	6	2.5	174	15 US-10-292-798-612	Sequence 612, App
68	6	2.5	178	15 US-10-292-798-606	Sequence 606, App
69	7	2.5	191	12 US-10-424-599-210718	Sequence 210718, A
70	7	2.5	194	9 US-09-864-761-33808	Sequence 33808, A
71	7	2.5	204	12 US-10-424-599-203852	Sequence 203852, A
72	7	2.5	220	12 US-10-282-122A-69434	Sequence 69434, A
73	7	2.5	228	15 US-10-264-049-4337	Sequence 4337, App
74	7	2.5	234	12 US-10-335-977-8107	Sequence 8107, App
75	7	2.5	235	12 US-10-335-977-8108	Sequence 8108, App
76	7	2.5	238	12 US-10-425-114-69447	Sequence 69447, A
77	7	2.5	247	12 US-10-450-055-38	Sequence 38, Appl
78	7	2.5	247	12 US-10-282-122A-54377	Sequence 54377, A
79	7	2.5	256	12 US-10-282-122A-67627	Sequence 67627, A
80	7	2.5	260	9 US-09-815-242-12110	Sequence 12110, A
81	7	2.5	260	12 US-10-282-122A-66844	Sequence 66844, A
82	7	2.5	261	12 US-10-282-122A-71312	Sequence 71312, A
83	7	2.5	262	12 US-10-282-122A-46629	Sequence 46629, A
84	7	2.5	262	12 US-10-282-122A-771319	Sequence 771319, A
85	7	2.5	265	12 US-10-282-122A-671132	Sequence 671132, A
86	7	2.5	268	12 US-10-282-122A-42798	Sequence 42798, A
87	7	2.5	268	12 US-10-282-122A-49256	Sequence 49256, A
88	7	2.5	268	12 US-10-282-122A-55568	Sequence 55568, A

89	2.5	268	12	US-10-282-122A-73574	Sequence 73574, A	162	7	2.5	475	14	US-10-299-642-24	Sequence 24, Appl
90	2.5	268	12	US-10-282-122A-74991	Sequence 74991, A	163	7	2.5	477	11	US-09-826-509-493	Sequence 493, Appl
91	2.5	268	12	US-10-282-122A-76001	Sequence 76001, A	164	7	2.5	477	12	US-09-964-956-47	Sequence 47, Appl
92	2.5	269	12	US-10-282-122A-58568	Sequence 58568, A	165	7	2.5	477	14	US-10-377-078-2	Sequence 2, Appl
93	2.5	269	12	US-10-282-122A-59439	Sequence 59439, A	166	7	2.5	477	14	US-10-377-078-3	Sequence 3, Appl
94	2.5	269	12	US-10-282-122A-68970	Sequence 68970, A	167	7	2.5	477	14	US-10-425-567A-100	Sequence 100, Appl
95	2.5	270	12	US-10-282-122A-45531	Sequence 45531, A	168	7	2.5	477	14	US-10-299-642-8	Sequence 8, Appl
96	2.5	271	9	US-09-815-242-11460	Sequence 11460, A	169	7	2.5	477	15	US-10-352-684A-10	Sequence 4, Appl
97	2.5	271	9	US-09-815-242-11624	Sequence 11624, A	170	7	2.5	480	12	US-09-964-956-15	Sequence 15, Appl
98	2.5	271	12	US-10-282-122A-59034	Sequence 59034, A	171	7	2.5	486	12	US-10-282-122A-70987	Sequence 70987, A
99	2.5	271	12	US-10-282-122A-71343	Sequence 71343, A	172	7	2.5	496	12	US-10-425-114-41748	Sequence 41748, A
100	2.5	271	12	US-10-282-122A-78174	Sequence 78174, A	173	7	2.5	496	12	US-10-282-122A-47105	Sequence 47105, A
101	2.5	271	12	US-10-282-122A-78367	Sequence 78367, A	174	7	2.5	507	9	US-09-963-791-10	Sequence 10, Appl
102	2.5	271	12	US-10-335-977-4882	Sequence 4882, Ap	175	7	2.5	507	9	US-10-419-276-10	Sequence 10, Appl
103	2.5	271	12	US-10-335-977-4883	Sequence 4883, Ap	176	7	2.5	556	14	US-10-332-585-7593	Sequence 7593, Ap
104	2.5	272	12	US-10-282-122A-52590	Sequence 52590, A	177	7	2.5	580	15	US-10-369-493-2839	Sequence 2839, Ap
105	2.5	274	12	US-10-282-122A-77956	Sequence 77956, A	178	7	2.5	589	9	US-09-963-791-12	Sequence 12, Appl
106	2.5	275	14	US-10-156-761-7735	Sequence 77356, A	179	7	2.5	589	9	US-10-419-276-12	Sequence 12, Appl
107	2.5	276	12	US-10-282-122A-60923	Sequence 60923, A	180	7	2.5	591	12	US-10-425-114-63542	Sequence 63542, A
108	2.5	280	12	US-10-335-977-4884	Sequence 4884, Ap	181	7	2.5	633	12	US-10-282-122A-34071	Sequence 34071, A
109	2.5	284	15	US-10-369-493-9042	Sequence 9042, Ap	182	7	2.5	652	9	US-09-815-242-13673	Sequence 13673, A
110	2.5	285	12	US-10-425-114-57973	Sequence 57973, A	183	7	2.5	652	9	US-10-282-122A-73973	Sequence 73973, A
111	2.5	289	14	US-10-156-761-8987	Sequence 8987, Ap	184	7	2.5	652	12	US-10-282-122A-47105	Sequence 47105, A
112	2.5	298	9	US-09-811-094-32	Sequence 32, Appl	185	7	2.5	670	12	US-08-834-666A-12	Sequence 12, Appl
113	2.5	298	9	US-09-811-094-32	Sequence 32, Appl	186	7	2.5	691	8	US-08-834-666A-22	Sequence 22, Appl
114	2.5	298	9	US-09-810-644-32	Sequence 32, Appl	187	7	2.5	691	8	US-09-738-626-5979	Sequence 6979, Ap
115	2.5	298	9	US-09-810-644-33	Sequence 33, Appl	188	7	2.5	692	9	US-09-963-791-24	Sequence 24, Appl
116	2.5	298	9	US-09-185-904A-32	Sequence 32, Appl	189	7	2.5	757	9	US-10-282-122A-65901	Sequence 65901, A
117	2.5	298	9	US-09-185-904A-33	Sequence 33, Appl	190	7	2.5	757	12	US-10-419-276-24	Sequence 24, Appl
118	2.5	298	12	US-10-403-571-154	Sequence 154, App	191	7	2.5	852	12	US-10-282-122A-65071	Sequence 65071, A
119	2.5	298	16	US-10-408-765A-1531	Sequence 1531, Ap	192	7	2.5	852	12	US-10-282-122A-65071	Sequence 65071, A
120	2.5	298	16	US-10-408-765A-2398	Sequence 2398, Ap	193	7	2.5	852	12	US-10-282-122A-65071	Sequence 65071, A
121	2.5	299	12	US-10-424-599-280669	Sequence 280669, A	194	7	2.5	859	9	US-09-918-171A-5	Sequence 5, Appl
122	2.5	300	15	US-10-369-493-5088	Sequence 5088, Ap	195	7	2.5	860	9	US-09-738-626-464	Sequence 464, Ap
123	2.5	309	14	US-10-304-928-10	Sequence 10, Appl	196	7	2.5	860	11	US-09-981-151A-34	Sequence 34, Appl
124	2.5	310	9	US-09-963-791-8	Sequence 8, Appl	197	7	2.5	875	12	US-10-282-122A-60678	Sequence 60678, A
125	2.5	310	12	US-10-419-276-8	Sequence 8, Appl	198	7	2.5	908	9	US-09-963-791-2	Sequence 2, Appl
126	2.5	311	12	US-10-282-122A-65475	Sequence 65475, A	199	7	2.5	908	12	US-10-419-276-2	Sequence 2, Appl
127	2.5	311	12	US-10-282-122A-65779	Sequence 65779, A	200	7	2.5	959	9	US-09-788-043C-1	Sequence 1, Appl
128	2.5	312	15	US-10-369-493-13581	Sequence 13581, A	201	7	2.5	1036	12	US-10-282-122A-76570	Sequence 76570, A
129	2.5	312	15	US-10-369-493-18216	Sequence 18216, A	202	7	2.5	1120	16	US-10-467-042-6	Sequence 6, Appl
130	2.5	317	9	US-09-963-791-16	Sequence 16, Appl	203	7	2.5	1120	16	US-10-282-122A-69138	Sequence 38, Appl
131	2.5	317	12	US-10-419-276-16	Sequence 16, Appl	204	7	2.5	1130	15	US-10-369-493-5580	Sequence 5580, Ap
132	2.5	332	15	US-10-369-493-8019	Sequence 8019, Ap	205	7	2.5	1130	15	US-10-369-493-5581	Sequence 5581, Ap
133	2.5	337	10	US-09-769-734-29	Sequence 53, Appl	206	7	2.5	1175	14	US-10-332-585-7014	Sequence 7014, Ap
134	2.5	337	12	US-10-107-431-53	Sequence 53, Appl	207	7	2.5	1178	12	US-10-282-122A-5234	Sequence 5234, A
135	2.5	341	15	US-10-369-493-641	Sequence 641, App	208	7	2.5	1233	9	US-09-738-626-4312	Sequence 4312, Ap
136	2.5	347	9	US-09-977-577-3	Sequence 3, Appl	209	7	2.5	1282	10	US-09-940-168-21	Sequence 21, Appl
137	2.5	352	9	US-09-815-242-10053	Sequence 10053, A	210	7	2.5	1488	10	US-09-940-168-17	Sequence 17, Appl
138	2.5	355	14	US-10-017-161-694	Sequence 694, App	211	7	2.5	1509	10	US-09-940-168-23	Sequence 23, Appl
139	2.5	355	14	US-10-017-161-722	Sequence 722, App	212	7	2.5	1517	12	US-09-940-168-19	Sequence 19, Appl
140	2.5	355	15	US-10-292-798-632	Sequence 632, App	213	7	2.5	1517	12	US-09-940-168-11	Sequence 11, Appl
141	2.5	356	12	US-09-963-791-20	Sequence 20, Appl	214	7	2.5	1517	12	US-10-376-537-71	Sequence 71, Appl
142	2.5	356	12	US-10-419-276-20	Sequence 20, Appl	215	7	2.5	1517	12	US-10-376-537-72	Sequence 72, Appl
143	2.5	376	12	US-10-425-114-55210	Sequence 55210, A	216	7	2.5	1972	12	US-10-702-148-21	Sequence 21, Appl
144	2.5	380	9	US-10-092-771-6	Sequence 6, Appl	217	7	2.5	1972	12	US-10-376-537-21	Sequence 21, Appl
145	2.5	386	12	US-10-424-599-21898	Sequence 21898, A	218	7	2.5	2154	12	US-10-282-122A-66695	Sequence 66695, A
146	2.5	411	14	US-10-350-598-1	Sequence 1, Appl	219	7	2.5	2154	12	US-10-282-122A-66695	Sequence 66695, A
147	2.5	419	9	US-09-712-363-689	Sequence 289, App	220	7	2.5	4688	12	US-10-282-122A-76655	Sequence 76655, A
148	2.5	419	12	US-10-282-122A-62797	Sequence 62797, A	221	7	2.5	6196	10	US-10-282-122A-67793	Sequence 67793, A
149	2.5	424	9	US-10-282-122A-64940	Sequence 64940, A	222	7	2.5	6196	10	US-09-940-168-77	Sequence 77, Appl
150	2.5	429	9	US-09-738-626-4306	Sequence 4306, Ap	223	7	2.5	6196	10	US-10-239-131A-474	Sequence 474, App
151	2.5	438	9	US-09-963-791-32	Sequence 32, Appl	224	7	2.5	6196	10	US-10-425-114-63542	Sequence 63542, A
152	2.5	438	12	US-10-419-276-22	Sequence 22, Appl	225	7	2.5	6196	10	US-10-425-114-63542	Sequence 63542, A
153	2.5	457	12	US-09-964-956-42	Sequence 49, Appl	226	7	2.5	6196	10	US-10-425-114-63542	Sequence 63542, A
154	2.5	457	12	US-09-964-956-49	Sequence 51, Appl	227	7	2.5	6196	10	US-10-425-114-63542	Sequence 63542, A
155	2.5	460	9	US-09-964-956-51	Sequence 51, Appl	228	7	2.5	6196	10	US-10-425-114-63542	Sequence 63542, A
156	2.5	468	12	US-09-738-626-4223	Sequence 4223, Ap	229	7	2.5	6196	10	US-10-425-114-63542	Sequence 63542, A
157	2.5	468	9	US-09-963-791-6	Sequence 6, Appl	230	7	2.5	6196	10	US-10-425-114-63542	Sequence 63542, A
158	2.5	468	12	US-10-419-276-6	Sequence 6, Appl	231	7	2.5	6196	10	US-10-425-114-63542	Sequence 63542, A
159	2.5	468	15	US-10-369-493-6465	Sequence 6465, Ap	232	7	2.5	6196	10	US-10-425-114-63542	Sequence 63542, A
160	2.5	471	15	US-10-369-493-21077	Sequence 21077, A	233	7	2.5	6196	10	US-10-425-114-63542	Sequence 63542, A
161	2.5	475	12	US-09-964-956-48	Sequence 1456, Ap	234	7	2.5	6196	10	US-10-425-114-63542	Sequence 63542, A

235	6	2.2	29	16	US-10-253-493-1431	Sequence 1431, Ap	308	6	2.2	103	12	US-10-424-599-237711	Sequence 237711, Ap
236	6	2.2	31	13	US-10-032-330-38	Sequence 38, Appl	309	6	2.2	103	14	US-10-032-585-1785	Sequence 7785, Ap
237	6	2.2	35	13	US-10-016-634A-132	Sequence 132, Appl	310	6	2.2	105	12	US-10-424-599-338378	Sequence 283378, Ap
238	6	2.2	40	12	US-10-424-599-153611	Sequence 153611, Ap	311	6	2.2	105	12	US-10-437-421-11	Sequence 11, Appl
239	6	2.2	43	12	US-10-424-599-186934	Sequence 186934, Ap	312	6	2.2	105	15	US-10-074-421-140	Sequence 140, Ap
240	6	2.2	43	14	US-10-321-857-136	Sequence 136, Appl	313	6	2.2	107	12	US-10-424-599-153985	Sequence 253985, Ap
241	6	2.2	43	14	US-10-318-675-136	Sequence 136, Appl	314	6	2.2	109	12	US-10-424-599-121413	Sequence 121413, Ap
242	6	2.2	45	14	US-10-029-386-28062	Sequence 28062, A	315	6	2.2	111	9	US-09-738-626-3554	Sequence 3654, Ap
243	6	2.2	50	12	US-10-424-599-239203	Sequence 239203, Ap	316	6	2.2	111	11	US-09-864-408A-170	Sequence 170, Ap
244	6	2.2	51	12	US-10-424-599-182141	Sequence 182141, Ap	317	6	2.2	113	9	US-10-424-599-195737	Sequence 195737, Ap
245	6	2.2	52	10	US-09-764-872-261	Sequence 261, Appl	318	6	2.2	114	9	US-09-811-284-149	Sequence 149, Appl
246	6	2.2	52	12	US-10-424-599-226535	Sequence 226535, Ap	319	6	2.2	116	12	US-10-282-122A-73031	Sequence 73031, A
247	6	2.2	52	12	US-10-424-599-240019	Sequence 240019, Ap	320	6	2.2	116	12	US-10-424-599-198173	Sequence 198173, Ap
248	6	2.2	53	12	US-10-424-599-225432	Sequence 225432, Ap	321	6	2.2	117	11	US-09-981-151A-79	Sequence 79, Appl
249	6	2.2	54	9	US-09-939-980-320	Sequence 320, Appl	322	6	2.2	117	11	US-09-864-408A-8366	Sequence 8366, Ap
250	6	2.2	55	12	US-10-424-599-279996	Sequence 279996, Ap	323	6	2.2	117	12	US-10-424-599-170868	Sequence 170868, Ap
251	6	2.2	56	12	US-10-424-599-167091	Sequence 167091, Ap	324	6	2.2	119	9	US-09-858-081-4	Sequence 4, Appl
252	6	2.2	57	12	US-10-424-599-153661	Sequence 153661, Ap	325	6	2.2	119	9	US-09-858-068-4	Sequence 4, Appl
253	6	2.2	58	12	US-10-424-599-156664	Sequence 156664, Ap	326	6	2.2	119	12	US-09-961-656-7	Sequence 7, Appl
254	6	2.2	58	14	US-10-029-386-31026	Sequence 31026, A	327	6	2.2	119	13	US-10-014-070-7	Sequence 7, Appl
255	6	2.2	60	12	US-10-424-599-244607	Sequence 244607, Ap	328	6	2.2	119	13	US-10-153-316-4	Sequence 4, Appl
256	6	2.2	62	12	US-10-424-599-148564	Sequence 148564, Ap	329	6	2.2	119	14	US-10-103-377C-4	Sequence 4, Appl
257	6	2.2	63	15	US-10-435-766-78	Sequence 78, Appl	330	6	2.2	119	15	US-10-028-248A-85	Sequence 85, Appl
258	6	2.2	64	16	US-10-280-139-10	Sequence 10, Appl	331	6	2.2	119	15	US-10-107-782-85	Sequence 85, Appl
259	6	2.2	65	12	US-10-424-599-283916	Sequence 283916, Ap	332	6	2.2	120	12	US-10-425-114-62096	Sequence 25944, A
260	6	2.2	66	10	US-09-983-802-654	Sequence 654, Appl	333	6	2.2	120	12	US-10-424-599-174541	Sequence 174541, Ap
261	6	2.2	66	12	US-09-973-278-372	Sequence 372, Appl	334	6	2.2	121	12	US-10-424-599-175317	Sequence 175317, Ap
262	6	2.2	66	12	US-10-424-599-234063	Sequence 234063, Ap	335	6	2.2	121	12	US-10-424-599-175317	Sequence 175317, Ap
263	6	2.2	66	12	US-10-424-599-256782	Sequence 256782, Ap	336	6	2.2	121	12	US-10-425-114-1185	Sequence 1185, A
264	6	2.2	66	12	US-10-424-599-277312	Sequence 277312, Ap	337	6	2.2	121	15	US-10-264-237-2563	Sequence 2563, Ap
265	6	2.2	66	12	US-10-424-599-285501	Sequence 285501, Ap	338	6	2.2	123	9	US-09-864-761-18806	Sequence 48606, A
266	6	2.2	66	12	US-09-984-490-654	Sequence 654, Appl	339	6	2.2	123	9	US-09-864-761-18806	Sequence 48606, A
267	6	2.2	67	12	US-10-424-599-214538	Sequence 214538, Ap	340	6	2.2	123	12	US-10-424-599-142864	Sequence 142864, Ap
268	6	2.2	67	12	US-10-424-599-262099	Sequence 262099, Ap	341	6	2.2	123	12	US-10-424-599-178133	Sequence 178133, Ap
269	6	2.2	68	12	US-10-424-599-213493	Sequence 213493, Ap	342	6	2.2	124	12	US-10-424-599-244845	Sequence 244845, Ap
270	6	2.2	68	12	US-10-424-599-281467	Sequence 281467, Ap	343	6	2.2	126	12	US-10-424-599-247128	Sequence 247128, Ap
271	6	2.2	72	12	US-10-424-599-183635	Sequence 183635, Ap	344	6	2.2	127	12	US-10-424-599-215154	Sequence 215154, Ap
272	6	2.2	72	12	US-10-424-599-226659	Sequence 226659, Ap	345	6	2.2	127	14	US-10-156-761-11456	Sequence 11456, A
273	6	2.2	73	12	US-10-424-599-168202	Sequence 168202, Ap	346	6	2.2	128	12	US-10-282-122A-5118	Sequence 5118, A
274	6	2.2	74	14	US-10-424-599-283720	Sequence 283720, Ap	347	6	2.2	128	12	US-10-282-122A-5118	Sequence 5118, A
275	6	2.2	74	14	US-10-029-386-31524	Sequence 31524, A	348	6	2.2	129	12	US-10-424-599-226533	Sequence 226533, Ap
276	6	2.2	75	12	US-10-424-599-154202	Sequence 154202, Ap	349	6	2.2	131	10	US-09-993-190-175	Sequence 190, Appl
277	6	2.2	76	12	US-10-424-599-179447	Sequence 179447, Ap	350	6	2.2	131	16	US-10-461-990-30	Sequence 30, Appl
278	6	2.2	80	12	US-10-424-599-229412	Sequence 229412, Ap	351	6	2.2	132	11	US-09-864-408A-2114	Sequence 2114, Ap
279	6	2.2	81	12	US-10-282-122A-52594	Sequence 52594, A	352	6	2.2	132	15	US-10-424-599-240909	Sequence 240909, Ap
280	6	2.2	81	12	US-10-424-599-160522	Sequence 160522, Ap	353	6	2.2	132	15	US-10-094-886-20	Sequence 20, Appl
281	6	2.2	82	10	US-09-903-463-16	Sequence 16, Appl	354	6	2.2	133	14	US-10-195-730-140	Sequence 140, Appl
282	6	2.2	83	12	US-10-424-599-267319	Sequence 267319, Ap	355	6	2.2	135	11	US-09-833-245-170	Sequence 170, Appl
283	6	2.2	84	12	US-10-282-122A-60406	Sequence 60406, A	356	6	2.2	135	12	US-10-424-599-184282	Sequence 184282, Ap
284	6	2.2	84	12	US-10-282-122A-60406	Sequence 60406, A	357	6	2.2	135	12	US-10-424-599-276407	Sequence 276407, Ap
285	6	2.2	84	15	US-10-074-978A-291	Sequence 291, Appl	358	6	2.2	136	12	US-10-424-599-148636	Sequence 148636, Ap
286	6	2.2	85	12	US-10-424-599-164049	Sequence 164049, Ap	359	6	2.2	136	12	US-10-424-599-153334	Sequence 153334, Ap
287	6	2.2	85	12	US-10-424-599-267982	Sequence 267982, Ap	360	6	2.2	138	12	US-10-424-599-239007	Sequence 239007, Ap
288	6	2.2	85	12	US-10-424-599-280762	Sequence 280762, Ap	361	6	2.2	139	12	US-10-424-599-278923	Sequence 278923, Ap
289	6	2.2	86	12	US-10-424-599-143476	Sequence 143476, Ap	362	6	2.2	139	12	US-10-627-676-574	Sequence 574, Appl
290	6	2.2	87	9	US-09-764-864-1130	Sequence 1130, Ap	363	6	2.2	139	12	US-10-450-055-44	Sequence 44, Appl
291	6	2.2	87	12	US-10-424-599-283898	Sequence 283898, Ap	364	6	2.2	140	15	US-10-108-860A-2564	Sequence 2564, Ap
292	6	2.2	88	13	US-10-264-04-3645	Sequence 3645, Ap	365	6	2.2	141	15	US-10-425-114-68389	Sequence 68389, Ap
293	6	2.2	89	9	US-09-764-864-1545	Sequence 1545, Ap	366	6	2.2	141	15	US-10-074-978A-14	Sequence 14, Appl
294	6	2.2	89	12	US-10-424-599-171497	Sequence 171497, Ap	367	6	2.2	142	12	US-10-424-599-205163	Sequence 205163, Ap
295	6	2.2	89	12	US-10-424-599-173898	Sequence 173898, Ap	368	6	2.2	142	12	US-10-424-599-198889	Sequence 198889, Ap
296	6	2.2	90	12	US-10-424-599-185961	Sequence 185961, Ap	369	6	2.2	144	12	US-10-282-122A-76111	Sequence 76111, A
297	6	2.2	91	12	US-10-424-599-283528	Sequence 283528, Ap	370	6	2.2	144	12	US-10-282-122A-76111	Sequence 76111, A
298	6	2.2	93	12	US-10-424-599-180542	Sequence 180542, Ap	371	6	2.2	144	12	US-10-424-599-149499	Sequence 149499, Ap
299	6	2.2	93	12	US-10-424-599-211048	Sequence 211048, Ap	372	6	2.2	145	12	US-10-282-122A-59117	Sequence 59117, A
300	6	2.2	93	15	US-10-074-978A-142	Sequence 142, Appl	373	6	2.2	146	14	US-10-091-007-196	Sequence 196, Appl
301	6	2.2	93	15	US-10-424-599-279209	Sequence 279209, Ap	374	6	2.2	147	9	US-09-895-913A-346	Sequence 346, Appl
302	6	2.2	96	12	US-10-424-599-228371	Sequence 228371, Ap	375	6	2.2	147	12	US-10-282-122A-56453	Sequence 56453, A
303	6	2.2	97	12	US-10-424-599-265936	Sequence 265936, Ap	376	6	2.2	148	12	US-10-282-122A-60619	Sequence 60619, A
304	6	2.2	99	11	US-09-864-408A-1178	Sequence 1178, Ap	377	6	2.2	148	12	US-10-282-122A-67938	Sequence 67938, A
305	6	2.2	100	11	US-10-282-122A-63523	Sequence 63523, A	378	6	2.2	148	12	US-10-282-122A-65305	Sequence 65305, A
306	6	2.2	100	15	US-10-108-260A-4375	Sequence 4375, Ap	379	6	2.2	149	12	US-10-282-122A-42921	Sequence 42921, A
307	6	2.2	102	12	US-10-424-599-267532	Sequence 267532, Ap	380	6	2.2	149	12	US-10-424-599-207991	Sequence 207991, Ap

381	6	2.2	150	10	US-09-764-891-4480	Sequence 4480, Ap	454	6	2.2	166	13	US-10-051-643-109	Sequence 109, App
382	6	2.2	151	12	US-10-282-122A-62416	Sequence 62416, A	455	6	2.2	166	14	US-10-305-979-36	Sequence 36, App1
383	6	2.2	151	12	US-10-282-122A-64879	Sequence 64879, A	456	6	2.2	166	15	US-10-389-674-33	Sequence 38, App1
384	6	2.2	151	13	US-10-062-254-52	Sequence 52, App1	457	6	2.2	166	15	US-10-389-674-42	Sequence 42, App1
385	6	2.2	151	15	US-10-094-886-18	Sequence 18, App1	458	6	2.2	166	15	US-10-389-674-43	Sequence 43, App1
386	6	2.2	153	12	US-10-282-122A-63692	Sequence 63692, A	459	6	2.2	166	15	US-10-389-674-44	Sequence 44, App1
387	6	2.2	153	14	US-10-091-007-146	Sequence 146, App	460	6	2.2	166	15	US-10-389-674-46	Sequence 46, App1
388	6	2.2	154	9	US-09-726-643-77	Sequence 77, App1	461	6	2.2	166	15	US-10-389-674-48	Sequence 48, App1
389	6	2.2	154	12	US-10-016-248-36	Sequence 36, App1	462	6	2.2	166	15	US-10-389-674-50	Sequence 50, App1
390	6	2.2	154	13	US-10-042-141-77	Sequence 77, App1	463	6	2.2	166	15	US-10-389-674-51	Sequence 51, App1
391	6	2.2	154	14	US-10-220-289-2	Sequence 2, App1	464	6	2.2	166	15	US-10-389-674-54	Sequence 54, App1
392	6	2.2	154	14	US-10-156-761-9045	Sequence 9045, Ap	465	6	2.2	166	15	US-10-389-674-56	Sequence 56, App1
393	6	2.2	155	9	US-09-925-301-1561	Sequence 1561, Ap	466	6	2.2	166	15	US-10-389-674-57	Sequence 57, App1
394	6	2.2	155	9	US-09-738-973-114	Sequence 114, App	467	6	2.2	166	15	US-10-389-674-59	Sequence 59, App1
395	6	2.2	155	9	US-09-854-133-114	Sequence 114, App	468	6	2.2	166	15	US-10-389-674-60	Sequence 60, App1
396	6	2.2	155	12	US-10-335-977-8128	Sequence 8128, Ap	469	6	2.2	166	15	US-10-389-674-61	Sequence 61, App1
397	6	2.2	155	14	US-09-815-242-12062	Sequence 114, App	470	6	2.2	166	15	US-10-389-674-62	Sequence 62, App1
398	6	2.2	156	9	US-10-144-649A-114	Sequence 114, App	471	6	2.2	166	15	US-10-389-674-63	Sequence 63, App1
399	6	2.2	156	12	US-09-764-891-3400	Sequence 3400, Ap	472	6	2.2	166	15	US-10-389-674-64	Sequence 64, App1
400	6	2.2	156	12	US-10-282-122A-66739	Sequence 66739, A	473	6	2.2	166	15	US-10-389-674-66	Sequence 66, App1
401	6	2.2	156	12	US-10-424-599-168143	Sequence 168143, A	474	6	2.2	166	15	US-10-389-674-67	Sequence 67, App1
402	6	2.2	156	12	US-10-424-599-168143	Sequence 168143, A	475	6	2.2	166	15	US-10-389-674-69	Sequence 69, App1
403	6	2.2	157	14	US-10-156-761-9744	Sequence 9744, Ap	476	6	2.2	166	15	US-10-389-674-70	Sequence 70, App1
404	6	2.2	157	16	US-10-415-969-40	Sequence 40, App1	477	6	2.2	166	15	US-10-389-674-79	Sequence 79, App1
405	6	2.2	158	12	US-09-738-626-6824	Sequence 6824, Ap	478	6	2.2	166	15	US-10-389-674-80	Sequence 80, App1
406	6	2.2	158	12	US-10-424-599-178819	Sequence 178819, A	479	6	2.2	166	15	US-10-389-674-84	Sequence 84, App1
407	6	2.2	158	12	US-10-335-977-8129	Sequence 8129, Ap	480	6	2.2	166	15	US-10-615-723-9	Sequence 9, App1
408	6	2.2	159	10	US-09-862-540-47	Sequence 47, App1	481	6	2.2	166	15	US-10-615-723-11	Sequence 11, App1
409	6	2.2	159	12	US-10-425-114-5572	Sequence 5572, A	482	6	2.2	166	15	US-10-615-723-30	Sequence 30, App1
410	6	2.2	159	15	US-10-369-493-16130	Sequence 16130, A	483	6	2.2	166	15	US-10-615-723-32	Sequence 32, App1
411	6	2.2	160	12	US-10-424-599-23206	Sequence 23206, A	484	6	2.2	166	15	US-10-615-723-33	Sequence 33, App1
412	6	2.2	160	12	US-10-425-114-62350	Sequence 62350, A	485	6	2.2	166	15	US-10-615-723-36	Sequence 36, App1
413	6	2.2	160	12	US-10-425-114-62350	Sequence 62350, A	486	6	2.2	166	15	US-10-615-723-38	Sequence 38, App1
414	6	2.2	161	9	US-09-726-643-78	Sequence 78, App1	487	6	2.2	166	15	US-10-615-723-38	Sequence 38, App1
415	6	2.2	161	10	US-09-997-672-9	Sequence 9, App1	488	6	2.2	166	15	US-10-615-723-42	Sequence 42, App1
416	6	2.2	161	12	US-10-424-599-234753	Sequence 234753, A	489	6	2.2	166	15	US-10-615-723-42	Sequence 42, App1
417	6	2.2	161	13	US-10-042-141-78	Sequence 78, App1	490	6	2.2	166	15	US-10-415-969-30	Sequence 30, App1
418	6	2.2	162	12	US-10-424-599-275145	Sequence 275145, A	491	6	2.2	166	15	US-10-415-969-78	Sequence 78, App1
419	6	2.2	162	16	US-10-461-990-31	Sequence 31, App1	492	6	2.2	166	16	US-10-415-969-82	Sequence 82, App1
420	6	2.2	163	12	US-10-424-599-156728	Sequence 156728, A	493	6	2.2	166	16	US-10-415-969-82	Sequence 82, App1
421	6	2.2	164	9	US-09-950-313-1	Sequence 3, App1	494	6	2.2	166	16	US-10-415-969-84	Sequence 84, App1
422	6	2.2	164	9	US-09-950-313-3	Sequence 3, App1	495	6	2.2	166	16	US-10-415-969-86	Sequence 86, App1
423	6	2.2	165	12	US-10-282-122A-44521	Sequence 44521, A	496	6	2.2	167	12	US-10-282-122A-62656	Sequence 62656, A
424	6	2.2	165	14	US-10-106-698-6109	Sequence 6109, Ap	497	6	2.2	167	12	US-10-424-599-21925	Sequence 21925, A
425	6	2.2	166	9	US-09-559-671A-75	Sequence 75, App1	498	6	2.2	167	12	US-10-424-599-21927	Sequence 21927, A
426	6	2.2	166	9	US-09-559-671A-76	Sequence 76, App1	499	6	2.2	167	12	US-10-424-599-21927	Sequence 21927, A
427	6	2.2	166	9	US-09-559-671A-77	Sequence 77, App1	500	6	2.2	167	15	US-10-369-493-1845	Sequence 1845, Ap
428	6	2.2	166	9	US-09-559-671A-79	Sequence 79, App1	501	6	2.2	169	12	US-10-282-122A-57648	Sequence 57648, A
429	6	2.2	166	9	US-09-559-671A-81	Sequence 81, App1	502	6	2.2	169	12	US-10-424-599-232864	Sequence 232864, A
430	6	2.2	166	9	US-09-559-671A-82	Sequence 82, App1	503	6	2.2	169	15	US-10-424-599-232864	Sequence 232864, A
431	6	2.2	166	9	US-09-559-671A-84	Sequence 84, App1	504	6	2.2	170	9	US-09-989-919-78	Sequence 919, App
432	6	2.2	166	9	US-09-559-671A-85	Sequence 85, App1	505	6	2.2	170	12	US-10-424-599-189920	Sequence 189920, A
433	6	2.2	166	9	US-09-559-671A-86	Sequence 86, App1	506	6	2.2	171	12	US-10-424-599-215472	Sequence 215472, A
434	6	2.2	166	9	US-09-977-034-2	Sequence 2, App1	507	6	2.2	171	12	US-10-424-599-215472	Sequence 215472, A
435	6	2.2	166	9	US-09-977-034-10	Sequence 10, App1	508	6	2.2	172	12	US-10-424-599-177986	Sequence 177986, A
436	6	2.2	166	9	US-09-977-034-13	Sequence 13, App1	509	6	2.2	172	15	US-10-369-493-15378	Sequence 15378, A
437	6	2.2	166	9	US-09-977-034-14	Sequence 14, App1	510	6	2.2	172	15	US-10-369-493-15378	Sequence 15378, A
438	6	2.2	166	9	US-09-977-034-15	Sequence 15, App1	511	6	2.2	173	12	US-10-369-493-15746	Sequence 15746, A
439	6	2.2	166	9	US-09-977-034-17	Sequence 17, App1	512	6	2.2	174	12	US-10-424-599-176585	Sequence 176585, A
440	6	2.2	166	9	US-09-977-034-18	Sequence 18, App1	513	6	2.2	175	12	US-10-424-599-177986	Sequence 177986, A
441	6	2.2	166	9	US-09-977-034-19	Sequence 19, App1	514	6	2.2	176	14	US-10-127-032-122	Sequence 122, App
442	6	2.2	166	10	US-09-880-505-109	Sequence 109, App	515	6	2.2	176	14	US-10-424-599-144059	Sequence 144059, A
443	6	2.2	166	10	US-09-880-505-109	Sequence 109, App	516	6	2.2	179	12	US-10-424-599-144059	Sequence 144059, A
444	6	2.2	166	10	US-09-880-505-109	Sequence 109, App	517	6	2.2	179	12	US-10-424-599-144059	Sequence 144059, A
445	6	2.2	166	10	US-09-880-505-109	Sequence 109, App	518	6	2.2	179	12	US-10-424-599-144059	Sequence 144059, A
446	6	2.2	166	10	US-09-880-505-109	Sequence 109, App	519	6	2.2	180	12	US-10-424-599-240907	Sequence 240907, A
447	6	2.2	166	10	US-09-880-505-109	Sequence 109, App	520	6	2.2	180	12	US-10-424-599-240907	Sequence 240907, A
448	6	2.2	166	10	US-09-880-505-109	Sequence 109, App	521	6	2.2	181	12	US-10-424-599-240907	Sequence 240907, A
449	6	2.2	166	10	US-09-880-505-109	Sequence 109, App	522	6	2.2	182	9	US-09-791-171-12	Sequence 12, App1
450	6	2.2	166	10	US-09-880-505-109	Sequence 109, App	523	6	2.2	182	9	US-09-791-171-12	Sequence 12, App1
451	6	2.2	166	10	US-09-880-505-109	Sequence 109, App	524	6	2.2	182	10	US-09-791-171-12	Sequence 12, App1
452	6	2.2	166	10	US-09-880-505-109	Sequence 109, App	525	6	2.2	182	10	US-09-791-171-12	Sequence 12, App1
453	6	2.2	166	12	US-10-425-114-61680	Sequence 61680, A	526	6	2.2	182	12	US-10-424-599-185594	Sequence 185594, A

527	6	2.2	182	12	US-09-943-443-12	Sequence 12, Appl	600	6	2.2	210	12	US-10-425-114-51533	Sequence 51533, A
528	6	2.2	183	12	US-10-424-599-261099	Sequence 261099, A	601	6	2.2	212	12	US-10-282-122A-59410	Sequence 59410, A
529	6	2.2	183	12	US-10-424-599-285416	Sequence 285416, A	602	6	2.2	212	12	US-10-424-599-335192	Sequence 235192, A
530	6	2.2	185	12	US-10-425-114-45011	Sequence 45011, A	603	6	2.2	213	12	US-10-282-122A-44412	Sequence 44312, A
531	6	2.2	185	15	US-10-289-762-106	Sequence 106, Appl	604	6	2.2	213	12	US-10-424-599-145495	Sequence 145495, A
532	6	2.2	186	9	US-09-738-626-6231	Sequence 6231, Ap	605	6	2.2	213	14	US-10-291-190-48	Sequence 48, Appl
533	6	2.2	186	12	US-10-424-599-219603	Sequence 219603, Ap	606	6	2.2	214	14	US-10-032-585-7363	Sequence 7363, Ap
534	6	2.2	186	15	US-10-158-057-223	Sequence 223, App	607	6	2.2	215	12	US-10-424-599-205345	Sequence 205345, A
535	6	2.2	187	9	US-09-925-937-490	Sequence 490, App	608	6	2.2	216	12	US-10-282-122A-51000	Sequence 51000, A
536	6	2.2	188	12	US-10-424-599-215561	Sequence 215561, A	609	6	2.2	216	14	US-10-424-599-175265	Sequence 175265, A
537	6	2.2	188	12	US-10-425-114-695692	Sequence 695692, A	610	6	2.2	216	14	US-10-128-714-8153	Sequence 8153, Ap
538	6	2.2	189	9	US-09-919-497-73	Sequence 73, Appl	611	6	2.2	217	14	US-10-376-345-6	Sequence 4, Appl
539	6	2.2	189	10	US-09-962-625-2	Sequence 2, Appl	612	6	2.2	217	14	US-10-376-345-6	Sequence 6, Appl
540	6	2.2	189	12	US-10-424-599-189299	Sequence 189299, A	613	6	2.2	218	12	US-10-425-114-45449	Sequence 44449, A
541	6	2.2	189	12	US-09-881-050-18	Sequence 18, Appl	614	6	2.2	220	9	US-09-726-643-54	Sequence 54, Appl
542	6	2.2	189	12	US-09-881-050-19	Sequence 19, Appl	615	6	2.2	220	13	US-10-042-141-54	Sequence 54, Appl
543	6	2.2	189	12	US-09-881-050-22	Sequence 22, Appl	616	6	2.2	221	10	US-09-934-455-216	Sequence 216, App
544	6	2.2	189	12	US-09-881-050-23	Sequence 23, Appl	617	6	2.2	221	12	US-10-282-122A-63302	Sequence 63302, A
545	6	2.2	189	12	US-09-881-050-25	Sequence 25, Appl	618	6	2.2	221	12	US-10-412-6998-1862	Sequence 1862, App
546	6	2.2	189	12	US-09-881-050-28	Sequence 28, Appl	619	6	2.2	221	12	US-10-225-066A-998	Sequence 998, App
547	6	2.2	189	14	US-10-284-740-10	Sequence 10, Appl	620	6	2.2	221	15	US-10-369-493-13371	Sequence 13371, A
548	6	2.2	189	16	US-10-415-963-38	Sequence 38, Appl	621	6	2.2	223	15	US-10-374-80A-274	Sequence 274, App
549	6	2.2	189	16	US-10-415-963-42	Sequence 42, Appl	622	6	2.2	223	12	US-10-424-599-147557	Sequence 147557, A
550	6	2.2	189	16	US-10-415-963-44	Sequence 44, Appl	623	6	2.2	223	12	US-10-276-774-2617	Sequence 2617, Ap
551	6	2.2	189	16	US-10-415-963-46	Sequence 46, Appl	624	6	2.2	223	14	US-10-156-761-14920	Sequence 14920, A
552	6	2.2	189	16	US-10-415-963-50	Sequence 50, Appl	625	6	2.2	224	9	US-09-393-634-9	Sequence 9, Appl
553	6	2.2	189	16	US-10-415-963-52	Sequence 52, Appl	626	6	2.2	224	12	US-10-364-861-9	Sequence 9, Appl
554	6	2.2	189	16	US-10-415-963-54	Sequence 54, Appl	627	6	2.2	224	14	US-10-383-882-9	Sequence 9, Appl
555	6	2.2	189	16	US-10-415-963-56	Sequence 56, Appl	628	6	2.2	226	9	US-09-815-242-12455	Sequence 12455, A
556	6	2.2	189	16	US-10-415-963-62	Sequence 62, Appl	629	6	2.2	226	9	US-09-815-242-12783	Sequence 12783, A
557	6	2.2	189	16	US-10-415-963-64	Sequence 64, Appl	630	6	2.2	226	15	US-10-369-493-8950	Sequence 8950, Ap
558	6	2.2	189	16	US-10-415-963-66	Sequence 66, Appl	631	6	2.2	227	15	US-10-346-000A-16	Sequence 16, Appl
559	6	2.2	189	16	US-10-415-963-68	Sequence 68, Appl	632	6	2.2	228	12	US-10-282-122A-50077	Sequence 50077, A
560	6	2.2	189	16	US-10-415-963-70	Sequence 70, Appl	633	6	2.2	229	12	US-10-335-977-7858	Sequence 7858, Ap
561	6	2.2	189	16	US-10-415-963-72	Sequence 72, Appl	634	6	2.2	229	15	US-10-369-493-6851	Sequence 6851, Ap
562	6	2.2	190	9	US-09-764-864-1071	Sequence 1071, Ap	635	6	2.2	230	9	US-09-738-626-5703	Sequence 5703, Ap
563	6	2.2	190	12	US-10-425-114-50711	Sequence 50711, A	636	6	2.2	231	14	US-10-106-698-6212	Sequence 6212, Ap
564	6	2.2	190	12	US-10-425-114-54773	Sequence 54773, A	637	6	2.2	231	16	US-10-107-857-48	Sequence 48, Appl
565	6	2.2	193	14	US-10-156-761-11378	Sequence 11378, A	638	6	2.2	231	14	US-10-389-566-933	Sequence 933, App
566	6	2.2	193	15	US-10-369-493-22176	Sequence 22176, A	639	6	2.2	232	12	US-10-424-599-14531	Sequence 14531, A
567	6	2.2	195	9	US-09-864-761-47064	Sequence 47064, A	640	6	2.2	232	12	US-10-282-122A-74630	Sequence 74630, A
568	6	2.2	196	12	US-10-424-599-175546	Sequence 175546, A	641	6	2.2	234	12	US-10-424-599-146518	Sequence 146518, A
569	6	2.2	196	12	US-10-424-599-175546	Sequence 175546, A	642	6	2.2	235	12	US-10-282-122A-67332	Sequence 67332, A
570	6	2.2	197	12	US-10-282-122A-71273	Sequence 71273, A	643	6	2.2	235	12	US-10-425-114-59338	Sequence 59338, A
571	6	2.2	198	9	US-09-925-302-500	Sequence 500, App	644	6	2.2	236	9	US-09-738-626-6304	Sequence 6304, Ap
572	6	2.2	198	12	US-09-925-302-500	Sequence 500, App	645	6	2.2	236	12	US-10-424-599-250081	Sequence 250081, A
573	6	2.2	199	15	US-10-369-493-17664	Sequence 17664, App	646	6	2.2	237	12	US-10-282-122A-77310	Sequence 77310, A
574	6	2.2	200	9	US-09-815-242-11021	Sequence 11021, A	647	6	2.2	237	12	US-10-425-114-41334	Sequence 41334, A
575	6	2.2	200	9	US-09-738-626-6458	Sequence 6458, Ap	648	6	2.2	237	12	US-10-425-114-56314	Sequence 56314, A
576	6	2.2	200	12	US-10-282-122A-52322	Sequence 52322, A	649	6	2.2	238	15	US-10-104-047-38902	Sequence 38902, Ap
577	6	2.2	200	12	US-10-282-122A-58168	Sequence 58168, A	650	6	2.2	239	12	US-10-424-599-229052	Sequence 229052, A
578	6	2.2	200	14	US-10-156-761-12379	Sequence 12379, A	651	6	2.2	239	14	US-10-424-599-282440	Sequence 282440, A
579	6	2.2	201	12	US-10-282-122A-62266	Sequence 62266, A	652	6	2.2	239	12	US-10-219-220-65	Sequence 65, Appl
580	6	2.2	201	14	US-10-128-917-3153	Sequence 3153, Ap	653	6	2.2	239	14	US-10-156-761-8154	Sequence 8154, Ap
581	6	2.2	202	12	US-10-335-977-8986	Sequence 8986, Ap	654	6	2.2	239	15	US-10-363-840-54	Sequence 84, Appl
582	6	2.2	203	8	US-08-781-986A-5207	Sequence 5207, Ap	655	6	2.2	239	15	US-10-363-840-54	Sequence 84, Appl
583	6	2.2	203	12	US-10-329-624-5207	Sequence 5207, Ap	656	6	2.2	240	15	US-10-369-493-23445	Sequence 23445, A
584	6	2.2	204	12	US-10-424-599-144986	Sequence 144986, A	657	6	2.2	242	14	US-10-156-761-8902	Sequence 8902, Ap
585	6	2.2	204	12	US-10-424-599-150135	Sequence 150135, A	658	6	2.2	243	9	US-09-738-626-4852	Sequence 4852, Ap
586	6	2.2	205	12	US-10-282-122A-67633	Sequence 67633, A	659	6	2.2	243	9	US-09-738-626-6413	Sequence 6413, Ap
587	6	2.2	205	12	US-10-424-599-248233	Sequence 248233, A	660	6	2.2	243	12	US-10-424-599-261883	Sequence 261883, A
588	6	2.2	205	12	US-10-425-114-63848	Sequence 63848, A	661	6	2.2	243	12	US-10-674-476-368	Sequence 368, App
589	6	2.2	207	12	US-10-282-122A-45062	Sequence 45062, A	662	6	2.2	243	12	US-10-344-980-2	Sequence 2, Appl
590	6	2.2	208	9	US-09-738-626-6261	Sequence 6261, Ap	663	6	2.2	244	15	US-10-369-493-16542	Sequence 16542, A
591	6	2.2	208	12	US-10-424-599-193167	Sequence 193167, A	664	6	2.2	245	10	US-09-903-463-15	Sequence 15, Appl
592	6	2.2	209	9	US-09-815-242-5583	Sequence 5583, Ap	665	6	2.2	245	12	US-10-425-114-43245	Sequence 43245, A
593	6	2.2	209	9	US-09-815-242-11967	Sequence 11967, A	666	6	2.2	246	12	US-10-282-122A-44407	Sequence 44407, A
594	6	2.2	209	9	US-09-738-626-5728	Sequence 5728, Ap	667	6	2.2	247	12	US-10-424-599-155603	Sequence 155603, A
595	6	2.2	209	12	US-10-282-122A-66579	Sequence 66579, A	668	6	2.2	248	12	US-10-282-122A-57202	Sequence 57202, A
596	6	2.2	209	12	US-10-424-599-239311	Sequence 239311, A	669	6	2.2	248	14	US-10-156-761-11171	Sequence 11171, A
597	6	2.2	209	14	US-10-307-294-7	Sequence 7, Appl	670	6	2.2	248	15	US-10-369-493-6091	Sequence 6091, Ap
598	6	2.2	210	10	US-09-975-719-172	Sequence 172, App	671	6	2.2	249	9	US-09-984-186-18	Sequence 18, Appl
599	6	2.2	210	12	US-10-282-122A-55523	Sequence 55523, A	672	6	2.2	249	14	US-10-237-667-18	Sequence 18, Appl

673	6	2.2	249	14	US-10-237-708-18	Sequence 18, Appl	746	6	2.2	277	12	US-10-206-576-30	Sequence 30, Appl
674	6	2.2	249	14	US-10-237-866-18	Sequence 18, Appl	747	6	2.2	278	15	US-10-369-493-8275	Sequence 8275, Ap
675	6	2.2	249	14	US-10-237-871-18	Sequence 18, Appl	748	6	2.2	279	15	US-10-282-1228-63865	Sequence 63865, A
676	6	2.2	249	14	US-10-237-624-18	Sequence 18, Appl	749	6	2.2	279	14	US-10-156-761-14320	Sequence 14320, A
677	6	2.2	249	14	US-10-156-761-12944	Sequence 12944, A	750	6	2.2	280	8	US-08-781-9864-5195	Sequence 5195, Ap
678	6	2.2	249	16	US-10-702-536-18	Sequence 18, Appl	751	6	2.2	280	12	US-10-424-599-252283	Sequence 252283, A
679	6	2.2	249	16	US-10-702-636-18	Sequence 18, Appl	752	6	2.2	280	12	US-10-389-624-5195	Sequence 5195, Ap
680	6	2.2	250	10	US-09-769-736-98	Sequence 98, Appl	753	6	2.2	280	16	US-10-408-7654-151	Sequence 151, Appl
681	6	2.2	250	12	US-10-424-599-2179815	Sequence 179815, Sequence 284325,	754	6	2.2	281	14	US-10-376-435-52	Sequence 54, Appl
682	6	2.2	250	12	US-10-424-599-284325	Sequence 284325, Sequence 86, Appl	755	6	2.2	282	10	US-09-769-736-54	Sequence 1240, Ap
683	6	2.2	251	14	US-10-189-346-86	Sequence 86, Appl	756	6	2.2	282	16	US-10-389-566-1240	Sequence 1273, A
684	6	2.2	251	15	US-10-910-154-707	Sequence 707, App	757	6	2.2	283	14	US-09-815-242-131731	Sequence 11773, A
685	6	2.2	252	12	US-10-282-1228-70252	Sequence 7052, A	758	6	2.2	284	9	US-09-815-242-131731	Sequence 73708, A
686	6	2.2	252	12	US-10-424-599-180645	Sequence 180645, Sequence 272981,	759	6	2.2	284	12	US-10-282-1228-73708	Sequence 446, App
687	6	2.2	253	12	US-10-424-599-272981	Sequence 272981, Sequence 69050, A	760	6	2.2	284	12	US-10-627-776-446	Sequence 450, App
688	6	2.2	254	12	US-10-282-1228-68050	Sequence 69050, A	761	6	2.2	284	12	US-10-627-776-446	Sequence 10979, A
689	6	2.2	254	12	US-10-425-114-68110	Sequence 68110, A	762	6	2.2	284	14	US-10-156-761-10979	Sequence 4731, Ap
690	6	2.2	255	15	US-10-369-493-21110	Sequence 21110, A	763	6	2.2	284	15	US-10-369-493-7491	Sequence 7491, Ap
691	6	2.2	257	12	US-10-206-576-32	Sequence 32, Appl	764	6	2.2	285	15	US-10-369-493-8027	Sequence 8027, Ap
692	6	2.2	257	9	US-09-864-761-37821	Sequence 37821, A	765	6	2.2	285	15	US-10-369-493-13392	Sequence 13392, A
693	6	2.2	258	12	US-10-425-114-52260	Sequence 52260, A	766	6	2.2	285	15	US-10-369-493-13392	Sequence 146, App
694	6	2.2	258	12	US-10-425-114-62204	Sequence 62204, A	767	6	2.2	285	16	US-10-432-443-146	Sequence 62914, A
695	6	2.2	259	9	US-09-815-242-5146	Sequence 5146, Ap	768	6	2.2	286	12	US-10-282-1228-62914	Sequence 65725, A
696	6	2.2	259	12	US-10-282-1228-43469	Sequence 43469, A	769	6	2.2	287	12	US-10-282-1228-65725	Sequence 78036, A
697	6	2.2	259	12	US-10-282-1228-63628	Sequence 63628, A	770	6	2.2	287	12	US-10-282-1228-78036	Sequence 49435, A
698	6	2.2	259	16	US-10-432-443-133	Sequence 133, App	771	6	2.2	288	12	US-10-282-1228-49435	Sequence 65329, A
699	6	2.2	260	9	US-09-765-772-32	Sequence 32, Appl	772	6	2.2	288	12	US-10-282-1228-65329	Sequence 366, App
700	6	2.2	260	12	US-10-282-1228-45357	Sequence 45357, A	773	6	2.2	288	15	US-10-043-487-366	Sequence 366, App
701	6	2.2	260	12	US-10-282-1228-69360	Sequence 69360, A	774	6	2.2	288	15	US-10-369-493-9706	Sequence 9706, App
702	6	2.2	260	12	US-10-282-1228-70210	Sequence 70210, A	775	6	2.2	289	12	US-10-425-114-40612	Sequence 40612, A
703	6	2.2	260	12	US-10-424-599-274453	Sequence 274453, A	776	6	2.2	289	14	US-10-156-761-13759	Sequence 13759, A
704	6	2.2	261	12	US-10-282-1228-67692	Sequence 67692, A	777	6	2.2	289	15	US-10-369-493-18230	Sequence 18230, A
705	6	2.2	261	14	US-10-161-051-23	Sequence 23, Appl	778	6	2.2	290	12	US-10-282-1228-153686	Sequence 53686, A
706	6	2.2	261	15	US-10-369-493-7775	Sequence 7775, Ap	779	6	2.2	291	12	US-10-424-599-204055	Sequence 204055, Sequence 205948,
707	6	2.2	261	15	US-10-264-213-227	Sequence 227, App	780	6	2.2	291	12	US-10-424-599-207605	Sequence 207605, Sequence 520948,
708	6	2.2	262	12	US-10-282-1228-54537	Sequence 54537, A	781	6	2.2	292	12	US-10-156-761-14678	Sequence 14678, A
709	6	2.2	262	14	US-10-156-761-10500	Sequence 10500, A	782	6	2.2	292	14	US-10-335-977-5290	Sequence 5290, Ap
710	6	2.2	263	12	US-10-282-1228-48062	Sequence 48062, A	783	6	2.2	293	15	US-10-335-977-5290	Sequence 23670, A
711	6	2.2	264	12	US-10-282-1228-71721	Sequence 71721, A	784	6	2.2	293	15	US-10-369-493-23620	Sequence 1246, Ap
712	6	2.2	265	12	US-10-424-599-258008	Sequence 258008, Sequence 5161, Ap	785	6	2.2	294	16	US-10-389-566-1246	Sequence 1246, Ap
713	6	2.2	265	12	US-10-335-977-5161	Sequence 5161, Ap	786	6	2.2	294	16	US-10-389-566-1606	Sequence 1606, Ap
714	6	2.2	265	12	US-10-335-977-5332	Sequence 5332, Ap	787	6	2.2	295	16	US-10-389-566-645	Sequence 645, App
715	6	2.2	265	12	US-10-335-977-5333	Sequence 5333, Ap	788	6	2.2	296	9	US-09-815-242-1246	Sequence 5446, App
716	6	2.2	266	12	US-10-335-977-5160	Sequence 5160, Ap	789	6	2.2	296	14	US-09-789-0544-10	Sequence 10, Appl
717	6	2.2	266	12	US-10-156-761-15054	Sequence 15054, A	790	6	2.2	296	14	US-10-029-186-32377	Sequence 32377, A
718	6	2.2	266	15	US-10-320-797-3067	Sequence 3067, Ap	791	6	2.2	296	14	US-10-025-806-116	Sequence 116, App
719	6	2.2	267	12	US-10-282-1228-53361	Sequence 53361, A	792	6	2.2	296	15	US-10-374-780-824	Sequence 824, App
720	6	2.2	267	12	US-10-424-599-273056	Sequence 273056, Sequence 11314, A	793	6	2.2	297	12	US-10-425-114-58061	Sequence 58061, A
721	6	2.2	268	9	US-09-815-242-11314	Sequence 11314, A	794	6	2.2	297	15	US-10-369-493-43061	Sequence 43061, Ap
722	6	2.2	268	9	US-09-815-242-11495	Sequence 11495, A	795	6	2.2	298	12	US-10-282-1228-57061	Sequence 57061, A
723	6	2.2	268	10	US-09-882-227-533	Sequence 533, App	796	6	2.2	298	12	US-10-424-599-187636	Sequence 187636, Sequence 4, Appl
724	6	2.2	268	12	US-10-282-1228-58680	Sequence 58680, A	797	6	2.2	298	12	US-09-815-694-4	Sequence 6885, Ap
725	6	2.2	268	12	US-10-335-977-5359	Sequence 5359, Ap	798	6	2.2	298	15	US-10-369-493-6885	Sequence 6885, Ap
726	6	2.2	270	12	US-10-282-1228-46819	Sequence 46819, A	799	6	2.2	298	16	US-10-389-566-550	Sequence 550, App
727	6	2.2	270	12	US-10-282-1228-70899	Sequence 70899, A	800	6	2.2	298	16	US-10-389-566-644	Sequence 644, App
728	6	2.2	271	12	US-10-282-1228-71810	Sequence 71810, A	801	6	2.2	299	9	US-09-738-626-1209	Sequence 4209, Ap
729	6	2.2	271	12	US-10-335-977-5334	Sequence 5334, Ap	802	6	2.2	300	12	US-10-425-114-59605	Sequence 59605, A
730	6	2.2	272	12	US-10-282-1228-50375	Sequence 50375, A	803	6	2.2	300	15	US-10-369-493-21433	Sequence 21433, A
731	6	2.2	272	12	US-10-282-1228-51605	Sequence 51605, A	804	6	2.2	300	15	US-10-369-493-23126	Sequence 23126, A
732	6	2.2	273	15	US-10-369-493-12081	Sequence 12081, A	805	6	2.2	302	9	US-09-815-242-12336	Sequence 12336, A
733	6	2.2	274	12	US-10-282-1228-65474	Sequence 65474, A	806	6	2.2	302	9	US-09-815-242-12861	Sequence 12861, A
734	6	2.2	274	12	US-10-282-1228-65778	Sequence 65778, A	807	6	2.2	302	12	US-10-282-1228-44027	Sequence 44027, A
735	6	2.2	275	12	US-10-282-1228-66155	Sequence 66155, A	808	6	2.2	302	12	US-10-282-1228-70930	Sequence 70930, A
736	6	2.2	275	12	US-10-425-114-68522	Sequence 68522, A	809	6	2.2	302	12	US-10-424-599-275275	Sequence 275275, Sequence 16, Appl
737	6	2.2	275	14	US-10-295-220-26	Sequence 26, Appl	810	6	2.2	303	9	US-09-798-029-16	Sequence 16, Appl
738	6	2.2	275	15	US-10-369-493-13773	Sequence 13773, A	811	6	2.2	303	12	US-09-933-079-818	Sequence 818, App
739	6	2.2	276	15	US-10-369-493-13773	Sequence 13773, A	812	6	2.2	303	12	US-10-004-860-818	Sequence 818, App
740	6	2.2	277	8	US-08-834-705-24	Sequence 24, Appl	813	6	2.2	304	14	US-10-023-882-818	Sequence 818, App
741	6	2.2	277	12	US-10-282-1228-57209	Sequence 57209, A	814	6	2.2	304	14	US-10-425-114-59872	Sequence 59872, A
742	6	2.2	277	12	US-10-282-1228-58897	Sequence 58897, A	815	6	2.2	305	12	US-10-425-114-66119	Sequence 66119, A
743	6	2.2	277	12	US-10-424-599-160702	Sequence 160702, A	816	6	2.2	305	15	US-10-369-493-162	Sequence 162, App
744	6	2.2	277	12	US-10-425-114-43859	Sequence 43859, A	817	6	2.2	305	15	US-10-369-493-11533	Sequence 11533, A
745	6	2.2	277	12	US-10-335-977-5289	Sequence 5289, Ap	818	6	2.2	305	15	US-10-369-493-14443	Sequence 14443, A

819	6	2.2	305	15	US-10-369-493-14764	Sequence 14764, A	892	6	2.2	307	14	US-10-242-990-30	Sequence 30, Appl
820	6	2.2	305	15	US-10-369-493-15244	Sequence 15244, A	893	6	2.2	307	14	US-10-243-023-30	Sequence 30, Appl
821	6	2.2	306	10	US-09-769-736-112	Sequence 112, Appl	894	6	2.2	307	14	US-10-243-103-30	Sequence 30, Appl
822	6	2.2	306	12	US-09-758-759-115	Sequence 115, Appl	895	6	2.2	307	14	US-10-243-276-30	Sequence 30, Appl
823	6	2.2	306	15	US-10-369-493-7062	Sequence 7062, Ap	896	6	2.2	307	14	US-10-243-326-30	Sequence 30, Appl
824	6	2.2	307	12	US-10-282-122A-74940	Sequence 74940, A	897	6	2.2	307	14	US-10-243-364-30	Sequence 30, Appl
825	6	2.2	307	12	US-10-245-752-30	Sequence 30, Appl	898	6	2.2	307	14	US-10-243-494-30	Sequence 30, Appl
826	6	2.2	307	12	US-10-245-859-30	Sequence 30, Appl	899	6	2.2	307	14	US-10-244-995-30	Sequence 30, Appl
827	6	2.2	307	13	US-10-001-054-28	Sequence 28, Appl	900	6	2.2	307	14	US-10-245-230-30	Sequence 30, Appl
828	6	2.2	307	14	US-10-245-103-30	Sequence 30, Appl	901	6	2.2	307	14	US-10-245-253-30	Sequence 30, Appl
829	6	2.2	307	14	US-10-245-107-30	Sequence 30, Appl	902	6	2.2	307	14	US-10-245-479-30	Sequence 30, Appl
830	6	2.2	307	14	US-10-245-143-30	Sequence 30, Appl	903	6	2.2	307	14	US-10-245-499-30	Sequence 30, Appl
831	6	2.2	307	14	US-10-245-771-30	Sequence 30, Appl	904	6	2.2	307	14	US-10-245-772-30	Sequence 30, Appl
832	6	2.2	307	14	US-10-245-851-30	Sequence 30, Appl	905	6	2.2	307	14	US-10-245-811-30	Sequence 30, Appl
833	6	2.2	307	14	US-10-245-883-30	Sequence 30, Appl	906	6	2.2	307	14	US-10-245-812-30	Sequence 30, Appl
834	6	2.2	307	14	US-10-237-535-30	Sequence 30, Appl	907	6	2.2	307	14	US-10-245-852-30	Sequence 30, Appl
835	6	2.2	307	14	US-10-238-183-30	Sequence 30, Appl	908	6	2.2	307	14	US-10-245-875-30	Sequence 30, Appl
836	6	2.2	307	14	US-10-238-283-30	Sequence 30, Appl	909	6	2.2	307	14	US-10-245-881-30	Sequence 30, Appl
837	6	2.2	307	14	US-10-238-370-30	Sequence 30, Appl	910	6	2.2	307	14	US-10-245-911-30	Sequence 30, Appl
838	6	2.2	307	14	US-10-245-055-30	Sequence 30, Appl	911	6	2.2	307	14	US-10-245-913-30	Sequence 30, Appl
839	6	2.2	307	14	US-10-245-147-30	Sequence 30, Appl	912	6	2.2	307	14	US-10-246-080-30	Sequence 30, Appl
840	6	2.2	307	14	US-10-245-730-30	Sequence 30, Appl	913	6	2.2	307	14	US-10-246-121-30	Sequence 30, Appl
841	6	2.2	307	14	US-10-245-739-30	Sequence 30, Appl	914	6	2.2	307	14	US-10-246-305-30	Sequence 30, Appl
842	6	2.2	307	14	US-10-246-210-30	Sequence 30, Appl	915	6	2.2	307	14	US-10-246-329-30	Sequence 30, Appl
843	6	2.2	307	14	US-10-239-196-30	Sequence 30, Appl	916	6	2.2	307	14	US-10-247-036-30	Sequence 30, Appl
844	6	2.2	307	14	US-10-243-024-30	Sequence 30, Appl	917	6	2.2	307	14	US-10-243-255-30	Sequence 30, Appl
845	6	2.2	307	14	US-10-243-409-30	Sequence 30, Appl	918	6	2.2	307	14	US-10-245-810-30	Sequence 30, Appl
846	6	2.2	307	14	US-10-245-621-30	Sequence 30, Appl	919	6	2.2	307	14	US-10-245-910-30	Sequence 30, Appl
847	6	2.2	307	14	US-10-245-880-30	Sequence 30, Appl	920	6	2.2	307	14	US-10-246-098-30	Sequence 30, Appl
848	6	2.2	307	14	US-10-245-031-30	Sequence 30, Appl	921	6	2.2	307	14	US-10-237-496-30	Sequence 30, Appl
849	6	2.2	307	14	US-10-243-093-30	Sequence 30, Appl	922	6	2.2	307	14	US-10-242-074-30	Sequence 30, Appl
850	6	2.2	307	14	US-10-243-185-30	Sequence 30, Appl	923	6	2.2	307	14	US-10-242-505-30	Sequence 30, Appl
851	6	2.2	307	14	US-10-245-427-30	Sequence 30, Appl	924	6	2.2	307	14	US-10-242-574-30	Sequence 30, Appl
852	6	2.2	307	14	US-10-245-473-30	Sequence 30, Appl	925	6	2.2	307	14	US-10-243-261-30	Sequence 30, Appl
853	6	2.2	307	14	US-10-245-770-30	Sequence 30, Appl	926	6	2.2	307	14	US-10-243-282-30	Sequence 30, Appl
854	6	2.2	307	14	US-10-245-877-30	Sequence 30, Appl	927	6	2.2	307	14	US-10-243-402-30	Sequence 30, Appl
855	6	2.2	307	14	US-10-246-976-30	Sequence 30, Appl	928	6	2.2	307	14	US-10-243-431-30	Sequence 30, Appl
856	6	2.2	307	14	US-10-243-329-30	Sequence 30, Appl	929	6	2.2	307	14	US-10-244-164-30	Sequence 30, Appl
857	6	2.2	307	14	US-10-242-743-30	Sequence 30, Appl	930	6	2.2	307	14	US-10-244-972-30	Sequence 30, Appl
858	6	2.2	307	14	US-10-242-845-30	Sequence 30, Appl	931	6	2.2	307	14	US-10-197-942-30	Sequence 30, Appl
859	6	2.2	307	14	US-10-237-636-30	Sequence 30, Appl	932	6	2.2	307	14	US-10-238-196-30	Sequence 30, Appl
860	6	2.2	307	14	US-10-238-325-30	Sequence 30, Appl	933	6	2.2	307	14	US-10-245-013-30	Sequence 30, Appl
861	6	2.2	307	14	US-10-238-346-30	Sequence 30, Appl	934	6	2.2	307	15	US-10-264-213-150	Sequence 150, App
862	6	2.2	307	14	US-10-238-411-30	Sequence 30, Appl	935	6	2.2	308	9	US-09-815-242-13087	Sequence 13087, A
863	6	2.2	307	14	US-10-243-128-30	Sequence 30, Appl	936	6	2.2	308	12	US-10-282-122A-72664	Sequence 72664, A
864	6	2.2	307	14	US-10-243-425-30	Sequence 30, Appl	937	6	2.2	308	12	US-10-107-431-25	Sequence 25, Appl
865	6	2.2	307	14	US-10-243-474-30	Sequence 30, Appl	938	6	2.2	308	12	US-09-758-759-11	Sequence 11, Appl
866	6	2.2	307	14	US-10-245-876-30	Sequence 30, Appl	939	6	2.2	308	14	US-10-156-761-11436	Sequence 11436, A
867	6	2.2	307	14	US-10-242-653-30	Sequence 30, Appl	940	6	2.2	309	9	US-09-815-242-13711	Sequence 13711, A
868	6	2.2	307	14	US-10-243-167-30	Sequence 30, Appl	941	6	2.2	309	10	US-09-510-332-85	Sequence 85, Appl
869	6	2.2	307	14	US-10-243-389-30	Sequence 30, Appl	942	6	2.2	309	10	US-10-282-122A-45051	Sequence 45051, A
870	6	2.2	307	14	US-10-244-947-30	Sequence 30, Appl	943	6	2.2	309	12	US-10-282-122A-64247	Sequence 64247, A
871	6	2.2	307	14	US-10-244-966-30	Sequence 30, Appl	944	6	2.2	309	12	US-10-282-122A-74839	Sequence 74839, A
872	6	2.2	307	14	US-10-244-979-30	Sequence 30, Appl	945	6	2.2	309	12	US-10-282-122A-75392	Sequence 75392, A
873	6	2.2	307	14	US-10-245-070-30	Sequence 30, Appl	946	6	2.2	309	14	US-10-309-933-3	Sequence 3, Appl
874	6	2.2	307	14	US-10-245-127-30	Sequence 30, Appl	947	6	2.2	310	8	US-08-964-716-42	Sequence 42, Appl
875	6	2.2	307	14	US-10-245-207-30	Sequence 30, Appl	948	6	2.2	310	9	US-09-815-242-10667	Sequence 10667, A
876	6	2.2	307	14	US-10-245-646-30	Sequence 30, Appl	949	6	2.2	310	12	US-10-282-122A-72138	Sequence 72138, A
877	6	2.2	307	14	US-10-245-695-30	Sequence 30, Appl	950	6	2.2	310	15	US-10-369-493-1160	Sequence 1160, Ap
878	6	2.2	307	14	US-10-245-699-30	Sequence 30, Appl	951	6	2.2	311	12	US-10-282-122A-49082	Sequence 49082, A
879	6	2.2	307	14	US-10-245-737-30	Sequence 30, Appl	952	6	2.2	311	12	US-10-282-122A-63438	Sequence 63438, A
880	6	2.2	307	14	US-10-245-878-30	Sequence 30, Appl	953	6	2.2	311	12	US-10-424-599-261206	Sequence 261206, A
881	6	2.2	307	14	US-10-245-890-30	Sequence 30, Appl	954	6	2.2	311	14	US-10-156-761-10112	Sequence 10112, A
882	6	2.2	307	14	US-10-245-899-30	Sequence 30, Appl	955	6	2.2	312	12	US-10-282-122A-75790	Sequence 75790, A
883	6	2.2	307	14	US-10-245-900-30	Sequence 30, Appl	956	6	2.2	312	12	US-10-282-122A-75790	Sequence 75790, A
884	6	2.2	307	14	US-10-247-088-30	Sequence 30, Appl	957	6	2.2	315	12	US-10-425-114-52560	Sequence 52560, A
885	6	2.2	307	14	US-10-245-454-30	Sequence 30, Appl	958	6	2.2	315	14	US-10-332-585-7236	Sequence 7236, Ap
886	6	2.2	307	14	US-10-237-471-30	Sequence 30, Appl	959	6	2.2	316	12	US-10-282-122A-61028	Sequence 61028, A
887	6	2.2	307	14	US-10-238-324-30	Sequence 30, Appl	960	6	2.2	316	14	US-10-219-220-260	Sequence 220, App
888	6	2.2	307	14	US-10-238-324-30	Sequence 30, Appl	961	6	2.2	316	15	US-10-369-493-22263	Sequence 22263, A
889	6	2.2	307	14	US-10-241-860-30	Sequence 30, Appl	962	6	2.2	317	5	US-08-815-242-13654	Sequence 13654, A
890	6	2.2	307	14	US-10-242-173-30	Sequence 30, Appl	963	6	2.2	317	10	US-09-769-744A-116	Sequence 116, App
891	6	2.2	307	14	US-10-242-652-30	Sequence 30, Appl	964	6	2.2	317	10	US-09-769-744A-116	Sequence 116, App

PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 1410
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11067
LENGTH: 273
TYPE: PRT
ORGANISM: Haemophilus influenzae
US-09-815-242-11067

Query Match 6.9%; Score 19; DB 9; Length 273;
Best Local Similarity 100.0%; Pred. No. 6.9e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 LNNLVIVGNTFYVPLAGYS 125
Db 104 LNNLVIVGNTFYVPLAGYS 122

RESULT 3
US-10-282-122A-58240
Sequence 58240, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haseibeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: EITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 58240
LENGTH: 273
TYPE: PRT
ORGANISM: Haemophilus influenzae
US-10-282-122A-58240

Query Match 6.9%; Score 19; DB 12; Length 273;
Best Local Similarity 100.0%; Pred. No. 6.9e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 LNNLVIVGNTFYVPLAGYS 125
Db 104 LNNLVIVGNTFYVPLAGYS 122

RESULT 4
US-10-380-817-2
Sequence 2, Application US/10380817
Publication No. US20040039169A1
GENERAL INFORMATION:
APPLICANT: Thonnard, Joelle
TITLE OF INVENTION: HAEMOPHILUS INFLUENZAE BASB202
FILE REFERENCE: BM45419
CURRENT FILING DATE: 2003-03-18
PRIOR APPLICATION NUMBER: PCT/EP01/10979
PRIOR FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: GB 0022992.2
PRIOR FILING DATE: 2000-09-19
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 273
TYPE: PRT
ORGANISM: No. US20040039169A1-typeable Haemophilus influenzae
US-10-380-817-2

Query Match 6.9%; Score 19; DB 12; Length 273;
Best Local Similarity 100.0%; Pred. No. 6.9e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 LNNLVIVGNTFYVPLAGYS 125
Db 104 LNNLVIVGNTFYVPLAGYS 122

RESULT 5
US-10-380-817-4
Sequence 4, Application US/10380817
Publication No. US20040039169A1
GENERAL INFORMATION:
APPLICANT: Thonnard, Joelle
TITLE OF INVENTION: HAEMOPHILUS INFLUENZAE BASB202
FILE REFERENCE: BM45419
CURRENT FILING DATE: 2003-03-18
PRIOR APPLICATION NUMBER: PCT/EP01/10979
PRIOR FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: GB 0022992.2
PRIOR FILING DATE: 2000-09-19
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 273
TYPE: PRT
ORGANISM: No. US20040039169A1-typeable Haemophilus influenzae
US-10-380-817-4

Query Match 6.9%; Score 19; DB 12; Length 273;
Best Local Similarity 100.0%; Pred. No. 6.9e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 LNNLVIVGNTFYVPLAGYS 125
Db 104 LNNLVIVGNTFYVPLAGYS 122

RESULT 6
US-10-380-817-6
Sequence 6, Application US/10380817
Publication No. US20040039169A1

```
/ GENERAL INFORMATION:
/ APPLICANT: Thonnard, Joelle
/ TITLE OF INVENTION: HAEMOPHILUS INFLUENZAE BASB202
/ TITLE OF INVENTION: POLYPEPTIDE, PRODUCTION, VACCINE AND DIAGNOSTIC USE
/ FILE REFERENCE: BM45419
/ CURRENT APPLICATION NUMBER: US/10/380,817
/ PRIOR FILING DATE: 2003-03-18
/ PRIOR APPLICATION NUMBER: PCT/EP01/10979
/ PRIOR FILING DATE: 2001-09-18
/ PRIOR APPLICATION NUMBER: GB 0022992.2
/ NUMBER OF SEQ ID NOS: 19
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 6
/ LENGTH: 273
/ TYPE: PRT
/ ORGANISM: No. US20040039169A1-typeable Haemophilus influenzae
US-10-380-817-6

Query Match
Best Local Similarity 100.0%; Score 19; DB 12; Length 273;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 107 LNNLVIVGNFTVYPLAGYS 125
Db 104 LNNLVIVGNFTVYPLAGYS 122

RESULT 7
US-10-380-817-8
/ Sequence 8, Application US/10380817
/ Publication No. US20040039169A1
/ GENERAL INFORMATION:
/ APPLICANT: Thonnard, Joelle
/ TITLE OF INVENTION: HAEMOPHILUS INFLUENZAE BASB202
/ FILE REFERENCE: BM45419
/ CURRENT APPLICATION NUMBER: US/10/380,817
/ PRIOR FILING DATE: 2003-03-18
/ PRIOR APPLICATION NUMBER: PCT/EP01/10979
/ PRIOR FILING DATE: 2001-09-18
/ PRIOR APPLICATION NUMBER: GB 0022992.2
/ NUMBER OF SEQ ID NOS: 19
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 8
/ LENGTH: 273
/ TYPE: PRT
/ ORGANISM: No. US20040039169A1-typeable Haemophilus influenzae
US-10-380-817-8

Query Match
Best Local Similarity 100.0%; Score 19; DB 12; Length 273;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 107 LNNLVIVGNFTVYPLAGYS 125
Db 104 LNNLVIVGNFTVYPLAGYS 122

RESULT 8
US-10-380-817-10
/ Sequence 10, Application US/10380817
/ Publication No. US20040039169A1
/ GENERAL INFORMATION:
/ APPLICANT: Thonnard, Joelle
/ TITLE OF INVENTION: HAEMOPHILUS INFLUENZAE BASB202
/ TITLE OF INVENTION: POLYPEPTIDE, PRODUCTION, VACCINE AND DIAGNOSTIC USE
/ FILE REFERENCE: BM45419
/ CURRENT APPLICATION NUMBER: US/10/380,817
/ PRIOR FILING DATE: 2003-03-18
/ PRIOR APPLICATION NUMBER: PCT/EP01/10979
/ PRIOR FILING DATE: 2001-09-18
```

```
/ PRIOR APPLICATION NUMBER: GB 0022992.2
/ PRIOR FILING DATE: 2000-09-19
/ NUMBER OF SEQ ID NOS: 19
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 10
/ LENGTH: 273
/ TYPE: PRT
/ ORGANISM: No. US20040039169A1-typeable Haemophilus influenzae
US-10-380-817-10

Query Match
Best Local Similarity 100.0%; Score 19; DB 12; Length 273;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 107 LNNLVIVGNFTVYPLAGYS 125
Db 104 LNNLVIVGNFTVYPLAGYS 122

RESULT 9
US-10-282-122A-68391
/ Sequence 68391, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ PRIOR FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27/257,931
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PAM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 68391
/ LENGTH: 271
/ TYPE: PRT
/ ORGANISM: Proteus mirabilis
US-10-282-122A-68391

Query Match
Best Local Similarity 100.0%; Score 16; DB 12; Length 271;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 224 FVEKDSYNNIIVAR 239
Db 219 FVEKDSYNNIIVAR 234

RESULT 10

US-10-282-122A-52838
; Sequence 52838, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52838
; LENGTH: 270
; TYPE: PRT
; ORGANISM: Clostridium botulinum
US-10-282-122A-52838

Query Match 4.7%; Score 13; DB 12; Length 270;
Best Local Similarity 100.0%; Pred. No. 0.00046;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 134 LKDGATIAVNDP 146
Db 130 LKDGATIAVNDP 142

RESULT 11

US-10-282-122A-67387
; Sequence 67387, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 67387
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-10-282-122A-67387

Query Match 4.3%; Score 12; DB 12; Length 276;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 LLEKGLIKND 166
Db 155 LLEKGLIKND 166

RESULT 12

US-10-092-243A-15
; Sequence 15, Application US/10092243A
; Publication No. US20020197625A1
; GENERAL INFORMATION:
; APPLICANT: Hallman, Jeffrey D.
; TITLE OF INVENTION: Microbial Polynucleotides Expressed During Infection of
; FILE REFERENCE: MBH00-5051
; CURRENT FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: 60/147,551
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US00/21340
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Actinobacillus actinomycetemcomitans
US-10-092-243A-15

Query Match 4.0%; Score 11; DB 13; Length 67;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 38 TIKGVWAGPE 48
Db 19 TIKGVWAGPE 29

RESULT 13

US-10-282-122A-55826
; Sequence 55826, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
PRIOR FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55826
; LENGTH: 271
; TYPE: PRF
; ORGANISM: Enterobacter cloacae
US-10-282-122A-55826

Query Match 4.0%; Score 11; DB 12; Length 271;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 224 FVEDKDSPEYN 234
Db 219 FVEDKDSPEYN 229

RESULT 14
US-10-282-122A-76195
; Sequence 76195, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:

APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
PRIOR FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 76195
; LENGTH: 271
; TYPE: PRF
; ORGANISM: Salmonella typhi
US-10-282-122A-76195

Query Match 4.0%; Score 11; DB 12; Length 271;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 224 FVEDKDSPEYN 234
Db 219 FVEDKDSPEYN 229

RESULT 15

US-10-282-122A-44792
; Sequence 44792, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A

```

CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 44792
LENGTH: 276
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-10-282-122A-44792

Query Match
Best Local Similarity 100.0%; Score 10; DB 12; Length 276;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 195 ARAIDVDVLA 204
DB 190 ARAIDVDVLA 199

RESULT 16
US-10-282-122A-44788
Sequence 44788, Application US/10282122A
Publication No. US20040029128A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 44788
LENGTH: 256
TYPE: PRT
```

```

PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 44788
LENGTH: 241
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-10-282-122A-44788

Query Match
Best Local Similarity 100.0%; Score 9; DB 12; Length 241;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 148 NARAILLL 156
DB 109 NARAILLL 117

RESULT 17
US-10-282-122A-61127
Sequence 61127, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 61127
LENGTH: 256
TYPE: PRT
```



```
/ ORGANISM: Legionella pneumophila
/ FEATURE:
/ NAME/KEY: MISC FEATURE
/ LOCATION: (227)..(227)
/ OTHER INFORMATION: X=any amino acid
/ FEATURE:
/ NAME/KEY: MISC FEATURE
/ LOCATION: (241)..(241)
/ OTHER INFORMATION: X=any amino acid
/ FEATURE:
/ NAME/KEY: MISC FEATURE
/ LOCATION: (244)..(244)
/ OTHER INFORMATION: X=any amino acid
/ FEATURE:
/ NAME/KEY: MISC FEATURE
/ LOCATION: (248)..(248)
/ OTHER INFORMATION: X=any amino acid
US-10-282-122A-61127

Query Match      3.3%; Score 9; DB 12; Length 256;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      140 IAVNDPSN 148
      |||||
Db      125 IAVNDPSN 133

RESULT 18
US-10-282-122A-51371
/ Sequence 51371, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Treweek, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 51371
```

```
/ LENGTH: 262
/ TYPE: PRT
/ ORGANISM: Bordetella pertussis
US-10-282-122A-51371

Query Match      3.3%; Score 9; DB 12; Length 262;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      133 ELKDGATTA 141
      |||||
Db      119 ELKDGATTA 127

RESULT 19
US-10-282-122A-55804
/ Sequence 55804, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Treweek, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 55804
/ LENGTH: 270
/ TYPE: PRT
/ ORGANISM: Enterobacter cloacae
US-10-282-122A-55804

Query Match      3.3%; Score 9; DB 12; Length 270;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      176 DIVENPKL 184
      |||||
Db      165 DIVENPKL 173
```

```
RESULT 20
US-10-282-122A-60558
; Sequence 60558, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60558
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Listeria monocytogenes
US-10-282-122A-60558

Query Match          3.3%; Score 9, DB 12; Length 273;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      140 IAVNDPSN 148
Db      136 IAVNDPSN 144

RESULT 21
US-10-282-122A-77174
; Sequence 77174, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 77174
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Vibrio cholerae
US-10-282-122A-77174

Query Match          3.3%; Score 9, DB 12; Length 275;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      112 IVGNFVYP 120
Db      111 IVGNFVYP 119

RESULT 22
US-09-764-853-805
; Sequence 805, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P1206
; CURRENT APPLICATION NUMBER: US/09/764,853
; CURRENT FILING DATE: 2001-01-17
; Remaining Prior Application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 805
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (20)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (55)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (56)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
```

LOCATION: (74)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (90)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-853-805

Query Match 2.9%; Score 8; DB 9; Length 96;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 146 PSNLABAL 153
Db 76 PSNLABAL 83

RESULT 23
US-10-091-438-162
Sequence 162, Application US/10091438
Publication No. US2003007760661
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: P1217C1
CURRENT APPLICATION NUMBER: US/10/091,438
PRIOR FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 09/764, 879
PRIOR FILING DATE: 2001-01-17, 065
PRIOR APPLICATION NUMBER: 60/179, 065
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/180, 628
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/214, 886
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 60/217, 487
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225, 758
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220, 963
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/217, 496
PRIOR FILING DATE: 2000-07-11, 447
PRIOR APPLICATION NUMBER: 60/225, 447
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/218, 290
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/225, 757
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/226, 868
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/216, 647
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/225, 267
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/216, 880
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/225, 270
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/251, 869
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/235, 834
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/234, 274
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/234, 223
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/228, 924
PRIOR FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: 60/224, 518
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236, 369
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/224, 519

PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220, 964
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/241, 809
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/249, 299
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/236, 327
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/241, 785
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/244, 617
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 60/225, 268
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236, 368
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/251, 856
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/251, 868
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/229, 344
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/234, 997
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: 60/229, 343
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229, 345
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229, 287
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229, 513
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/231, 413
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/229, 509
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/236, 367
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/237, 039
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237, 038
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237, 040
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/240, 960
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/239, 935
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/239, 937
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/241, 787
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/246, 474
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/246, 532
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/249, 216
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249, 210
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/226, 681
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225, 759
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/225, 213
PRIOR FILING DATE: 2000-08-14

PRIOR APPLICATION NUMBER: 60/227,182
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225,214
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/235,836
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/230,438
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/215,135
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 60/225,266
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/249,218
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,208
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,213
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,212
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,207
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,245
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,244
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,217
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,211
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,215
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,264
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/232,400
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/231,242
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/232,081
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/232,080
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/231,414
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/231,244
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/233,064
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/233,063
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,397
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,399
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,401
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/241,808
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,826
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,786
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,221
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/246,475
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/231,243
PRIOR FILING DATE: 2000-09-08

Query Match 2.9%; Score 8; DB 14; Length 96;
Best Local Similarity 100.0%; Pred. No. 13;
Matches: 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 146 PSNLARL 153
DB 76 PSNLARL 83

RESULT 24
US-10-424-599-262926
Sequence 262926, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53223)B
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 262926
LENGTH: 143
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_79443C.1.pgp
US-10-424-599-262926

Query Match 2.9%; Score 8; DB 12; Length 143;
Best Local Similarity 100.0%; Pred. No. 18;
Matches: 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 188 EVDTSVAA 195
DB 93 EVDTSVAA 100

RESULT 25
US-09-815-242-11656
Sequence 11656, Application US/09815242
Patent No. US2002061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl U.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16


```
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282.122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 49339
LENGTH: 270
TYPE: PRF
ORGANISM: Burkholderia fungorum
US-10-282-122A-49339

Query Match          2.9%; Score 8; DB 12; Length 270;
Best Local Similarity 100.0%; Pred. No. 33;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      141 AVPNPDSN 148
DB      132 AVPNPDSN 139

RESULT 29
US-09-815-242-10050
Sequence 10050, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zykkind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815.242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 43291
LENGTH: 271
TYPE: PRF
ORGANISM: Escherichia coli
US-10-282-122A-43291
```

```
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10050
LENGTH: 271
TYPE: PRF
ORGANISM: Escherichia coli
US-09-815-242-10050

Query Match          2.9%; Score 8; DB 9; Length 271;
Best Local Similarity 100.0%; Pred. No. 33;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      113 VGNFTVYP 120
DB      106 VGNFTVYP 113

RESULT 30
US-10-282-122A-43291
Sequence 43291, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zykkind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282.122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 43291
LENGTH: 271
TYPE: PRF
ORGANISM: Escherichia coli
US-10-282-122A-43291
```



```

; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57911
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Enterococcus faecium
; US-10-282-122A-57911

```

```

Query Match      2.9%; Score 8; DB 12; Length 282;
Best Local Similarity 100.0%; Pred. No. 35;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      133 ELKDGTI 140
      |||||
Db      131 ELKDGTI 138

```

```

RESULT 34
US-10-282-122A-47704
; Sequence 47704, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578

```

```

; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47704
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Burkholderia cepacia
; US-10-282-122A-47704

```

```

Query Match      2.9%; Score 8; DB 12; Length 295;
Best Local Similarity 100.0%; Pred. No. 36;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      141 AVNDPSN 148
      |||||
Db      157 AVNDPSN 164

```

```

RESULT 35
US-09-764-853-621
; Sequence 621, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P206
; CURRENT APPLICATION NUMBER: US/09/764,853
; CURRENT FILING DATE: 2001-01-17
; Prior Application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 621
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-764-853-621

```

```

Query Match      2.9%; Score 8; DB 9; Length 379;
Best Local Similarity 100.0%; Pred. No. 46;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      146 PSNLARAL 153
      |||||
Db      128 PSNLARAL 135

```

```

RESULT 36
US-10-369-493-20865
; Sequence 20865, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374

```

SEQ ID NO 20865
TYPE: PRT
LENGTH: 467
ORGANISM: Rhodospseudomonas palustris
US-10-369-493-20865

Query Match
Best Local Similarity 100.0%; Pred. No. 56;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 135 KDGATIAV 142
DB 336 KDGATIAV 343

RESULT 37
US-10-369-493-15592
Sequence 15592, Application US/10369493
Publication No. US2003023367SAI
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 15592
LENGTH: 641
TYPE: PRT
ORGANISM: Xanthomonas campestris
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(641)
OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-15592

Query Match
Best Local Similarity 100.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 IALAGCSN 23
DB 2 IALAGCSN 9

RESULT 38
US-09-854-845-45
Sequence 45, Application US/09854845
Patent No. US20020098491A1
GENERAL INFORMATION:
APPLICANT: Walke, D. Wade
APPLICANT: Wang, Xiaoming
APPLICANT: Scoville, John
APPLICANT: Turner, C. Alexander Jr.
TITLE OF INVENTION: No. US20020098491A1 Human Semaphorin Homologs and Polynucleotide
FILE REFERENCE: LEX-0177-USA
CURRENT APPLICATION NUMBER: US/09/854,845
CURRENT FILING DATE: 2001-05-14
PRIOR APPLICATION NUMBER: US 60/205,274
PRIOR FILING DATE: 2000-05-18
PRIOR APPLICATION NUMBER: US 60/208,893
PRIOR FILING DATE: 2000-06-02
NUMBER OF SEQ ID NOS: 50
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 45
LENGTH: 739

TYPE: PRT
ORGANISM: homo sapiens
US-09-854-845-45

Query Match
Best Local Similarity 100.0%; Pred. No. 86;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 PSNLARAL 153
DB 488 PSNLARAL 495

RESULT 39
US-09-854-845-43
Sequence 43, Application US/09854845
Patent No. US20020098491A1
GENERAL INFORMATION:
APPLICANT: Walke, D. Wade
APPLICANT: Wang, Xiaoming
APPLICANT: Scoville, John
APPLICANT: Turner, C. Alexander Jr.
TITLE OF INVENTION: No. US20020098491A1 Human Semaphorin Homologs and Polynucleotide
FILE REFERENCE: LEX-0177-USA
CURRENT APPLICATION NUMBER: US/09/854,845
CURRENT FILING DATE: 2001-05-14
PRIOR APPLICATION NUMBER: US 60/205,274
PRIOR FILING DATE: 2000-05-18
PRIOR APPLICATION NUMBER: US 60/208,893
PRIOR FILING DATE: 2000-06-02
NUMBER OF SEQ ID NOS: 50
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 43
LENGTH: 744
TYPE: PRT
ORGANISM: homo sapiens
US-09-854-845-43

Query Match
Best Local Similarity 100.0%; Pred. No. 86;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 PSNLARAL 153
DB 493 PSNLARAL 500

RESULT 40
US-09-854-845-49
Sequence 49, Application US/09854845
Patent No. US20020098491A1
GENERAL INFORMATION:
APPLICANT: Walke, D. Wade
APPLICANT: Wang, Xiaoming
APPLICANT: Scoville, John
APPLICANT: Turner, C. Alexander Jr.
TITLE OF INVENTION: No. US20020098491A1 Human Semaphorin Homologs and Polynucleotide
FILE REFERENCE: LEX-0177-USA
CURRENT APPLICATION NUMBER: US/09/854,845
CURRENT FILING DATE: 2001-05-14
PRIOR APPLICATION NUMBER: US 60/205,274
PRIOR FILING DATE: 2000-05-18
PRIOR APPLICATION NUMBER: US 60/208,893
PRIOR FILING DATE: 2000-06-02
NUMBER OF SEQ ID NOS: 50
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 49
LENGTH: 766
TYPE: PRT
ORGANISM: homo sapiens
US-09-854-845-49

Query Match
2.9%; Score 8; DB 9; Length 766;

Tue Jun 22 11:04:02 2004

us-10-018-672-2.oligo.rapb

Page 23

Best Local Similarity 100.0%; Pred. No. 89;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 146 PSNLRAL 153
|||
Db 488 PSNLRAL 495

Search completed: June 16, 2004, 11:23:05
Job time : 59 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 16, 2004, 11:12:42 ; Search time 45 Seconds
(without alignments)
1935.178 Million cell updates/sec

Title: US-10-018-672-2
Perfect score: 276
Sequence: 1 MNFGKINGICALASGIALAG.....TDEVEAKKQKQFDGVIKGW 276

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1017041 seqs, 315518202 residues
Word size: 0

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

- 1: SPTRMBL.25.*
- 2: sp_archaea.*
- 3: sp_bacteria.*
- 4: sp_fungi.*
- 5: sp_human.*
- 6: sp_invertebrate.*
- 7: sp_mammal.*
- 8: sp_mhce.*
- 9: sp_organelle.*
- 10: sp_phage.*
- 11: sp_plant.*
- 12: sp_rodent.*
- 13: sp_virus.*
- 14: sp_vertebrate.*
- 15: sp_unclassified.*
- 16: sp_virus.*
- 17: sp_bacteriopl.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	5.8	270	16	Q7VM93 haemophilus
2	11	4.0	271	16	Q83MC6 shigella fl
3	10	3.6	271	16	Q7WIF5
4	10	3.6	271	16	Q7WYF6
5	10	3.6	281	16	Q7W6I3
6	10	3.3	259	2	O69441
7	9	3.3	262	16	Q7WEM2
8	9	3.3	262	16	Q7W3I7
9	9	3.3	262	16	Q7W3I7
10	9	3.3	273	16	Q8YA74
11	9	3.3	276	16	Q7WML8
12	9	3.3	276	16	Q7WBS1
13	9	3.3	276	16	Q7W271
14	9	3.3	278	16	Q8YD39
15	9	3.3	278	16	Q8FV86
16	8	2.9	89	17	Q9UYT7

17	8	2.9	255	2	Q9AGI3	Q9AGI3 pseudomonas
18	8	2.9	259	2	Q7XSM6	Q7XSM6 pseudomonas
19	8	2.9	259	16	Q8XNL7	Q8XNL7 clostridium
20	8	2.9	259	16	Q8U7G0	Q8U7G0 agrobacteri
21	8	2.9	265	16	Q7WV1	Q7WV1 bordetella
22	8	2.9	265	16	Q7W4E3	Q7W4E3 bordetella
23	8	2.9	265	16	Q7WV70	Q7WV70 bordetella
24	8	2.9	266	16	Q8P4S9	Q8P4S9 xanthomonas
25	8	2.9	269	16	Q7V195	Q7V195 helicobacte
26	8	2.9	269	16	Q8PGF0	Q8PGF0 xanthomonas
27	8	2.9	269	16	Q8WRS3	Q8WRS3 vibrio para
28	8	2.9	270	16	Q8CMC8	Q8CMC8 escherichia
29	8	2.9	272	16	Q8S5C6	Q8S5C6 clostridium
30	8	2.9	277	4	Q7Z3D8	Q7Z3D8 homo sapien
31	8	2.9	281	16	Q9A1E4	Q9A1E4 streptococc
32	8	2.9	281	16	Q8P2K7	Q8P2K7 streptococc
33	8	2.9	281	16	Q8XK89	Q8XK89 streptococc
34	8	2.9	305	3	Q9P876	Q9P876 pichia jadi
35	8	2.9	305	3	Q9P875	Q9P875 pichia jadi
36	8	2.9	347	17	Q8ZSM1	Q8ZSM1 pyrobaculum
37	8	2.9	351	16	Q8A673	Q8A673 bacteroides
38	8	2.9	374	2	Q93AL8	Q93AL8 clostridium
39	8	2.9	446	16	Q9HYE0	Q9HYE0 pseudomonas
40	8	2.9	512	16	Q89R04	Q89R04 bradyrhizob
41	8	2.9	556	17	Q8TX48	Q8TX48 methanopyru
42	8	2.9	652	11	Q8BRP3	Q8BRP3 mus musculu
43	8	2.9	652	16	Q8P550	Q8P550 xanthomonas
44	8	2.9	761	5	Q9V9X1	Q9V9X1 diosiphilia
45	8	2.9	1220	11	Q8K1G2	Q8K1G2 mus musculu
46	8	2.9	1406	5	Q8EAV3	Q8EAV3 dictyosteli
47	8	2.9	3665	16	Q7UC60	Q7UC60 rhodospirilli
48	8	2.5	27	5	Q9NUT6	Q9NUT6 leishmania
49	7	2.5	47	16	Q8F2S2	Q8F2S2 leptospira
50	7	2.5	68	6	Q23368	Q23368 sus scrofa
51	7	2.5	74	4	Q9NYS5	Q9NYS5 homo sapien
52	7	2.5	87	16	Q9PHX5	Q9PHX5 campylobact
53	7	2.5	92	13	Q7T209	Q7T209 creochromis
54	7	2.5	104	5	Q8TXK6	Q8TXK6 caenorhabdi
55	7	2.5	105	16	Q8COT9	Q8COT9 streptococc
56	7	2.5	109	13	Q7T217	Q7T217 creochromis
57	7	2.5	114	9	Q7Y272	Q7Y272 phage phi 4
58	7	2.5	114	16	Q8X4T0	Q8X4T0 escherichia
59	7	2.5	114	16	Q8X4K2	Q8X4K2 escherichia
60	7	2.5	114	16	Q8X2D2	Q8X2D2 escherichia
61	7	2.5	115	16	Q8XAF8	Q8XAF8 escherichia
62	7	2.5	115	16	Q8X8Y3	Q8X8Y3 escherichia
63	7	2.5	121	16	Q818H9	Q818H9 bacillus ce
64	7	2.5	123	16	Q8X321	Q8X321 escherichia
65	7	2.5	124	16	Q25416	Q25416 helicobacte
66	7	2.5	137	17	Q9HPK6	Q9HPK6 halobacteri
67	7	2.5	138	16	Q7V238	Q7V238 prochlorococ
68	7	2.5	142	16	Q8PNP6	Q8PNP6 xanthomonas
69	7	2.5	150	17	Q972M6	Q972M6 sulfobobus
70	7	2.5	151	8	Q472S0	Q472S0 streptococci
71	7	2.5	152	8	Q472B9	Q472B9 aristida co
72	7	2.5	154	16	Q55120	Q55120 synochocyst
73	7	2.5	156	6	Q28532	Q28532 unidentified
74	7	2.5	161	2	Q8VUO8	Q8VUO8 pseudomonas
75	7	2.5	161	16	Q8BP92	Q8BP92 pseudomonas
76	7	2.5	163	5	Q8ITU1	Q8ITU1 bathymodiol
77	7	2.5	163	5	Q8ITU1	Q8ITU1 bathymodiol
78	7	2.5	163	5	Q8ITU1	Q8ITU1 bathymodiol
79	7	2.5	170	16	Q8PC24	Q8PC24 xanthomonas
80	7	2.5	174	4	Q8NGW2	Q8NGW2 homo sapien
81	7	2.5	176	5	Q8NMW6	Q8NMW6 caenorhabdi
82	7	2.5	178	4	Q8NMW7	Q8NMW7 homo sapien
83	7	2.5	185	2	Q93A38	Q93A38 leptospira
84	7	2.5	186	5	Q9M587	Q9M587 diosiphilia
85	7	2.5	188	16	Q92RK4	Q92RK4 rhizobium m
86	7	2.5	188	16	Q8UH88	Q8UH88 agrobacteri
87	7	2.5	196	16	Q8YB34	Q8YB34 bruceella me
88	7	2.5	196	16	Q8FWZ9	Q8FWZ9 bruceella su
89	7	2.5	199	2	Q9R7F0	Q9R7F0 borrelia bu

90	2.5	207	3	Q9URD7	Q9urd7 aspergillus	13	7	2.5	300	16	Q8UDP3	Q8udp3 agrobacteri
91	2.5	207	16	Q8G9P3	Q8g9p3 rhizobium 1	14	7	2.5	300	17	Q8ZVU1	Q8zvu1 deinococcus
92	2.5	209	16	Q8KGA0	Q8kga0 chlorobium	15	7	2.5	300	17	Q971G8	Q971g8 sulfolobus
93	2.5	209	16	Q8GUK1	Q8guk1 bradyrhizob	16	7	2.5	303	2	Q93EAK	Q93eak rhizobium 1
94	2.5	215	10	Q49911	Q49911 nicotiana t	167	7	2.5	303	16	Q8CVA8	Q8cva8 oceanobacili
95	2.5	218	16	Q89120	Q89120 bradyrhizob	168	7	2.5	306	16	Q8U782	Q8u782 agrobacteri
96	2.5	220	16	Q889Q8	Q889q8 pseudomonas	169	7	2.5	307	8	Q9XM22	Q9xm22 ascaris suu
97	2.5	224	17	Q973Z5	Q973z5 sulfolobus	170	7	2.5	317	13	Q9CHT8	Q9cht8 lactococcus
98	2.5	231	16	Q83HV9	Q83hv9 tripheryma	171	7	2.5	322	16	Q89HG6	Q89hg6 bradyrhizob
99	2.5	231	16	Q83GCO	Q83gco nicotiana t	172	7	2.5	323	17	Q8EYX1	Q8eyx1 methanosarc
100	2.5	232	10	Q49910	Q49910 nicotiana t	173	7	2.5	335	17	Q8U752	Q8u752 methanosarc
101	2.5	232	10	Q9W3U3	Q9w3u3 nicotiana t	174	7	2.5	341	5	Q9NB46	Q9nb46 caenorhabdi
102	2.5	232	10	Q8ALC7	Q8alc7 chrysanthem	175	7	2.5	344	16	Q67697	Q67697 aquifex aeo
103	2.5	234	16	Q9ZLC9	Q9zlc9 halicobacte	176	7	2.5	350	16	Q87CL9	Q87cl9 xylella fas
104	2.5	239	16	Q9PMD0	Q9pmd0 campylobact	177	7	2.5	352	16	Q8Y4K6	Q8y4k6 escherichia
105	2.5	241	16	Q9ASP2	Q9asp2 caulobacter	178	7	2.5	355	4	Q8NGU3	Q8ngu3 homo sapien
106	2.5	242	16	Q83CE8	Q83ce8 bradyrhizob	179	7	2.5	358	17	Q8THV7	Q8thv7 borrelia
107	2.5	244	16	Q9BD84	Q9bd84 rhizobium 1	180	7	2.5	363	16	Q8PPU3	Q8ppv3 methanosarc
108	2.5	247	16	Q9PHP4	Q9php4 campylobact	181	7	2.5	363	16	Q8P501	Q8p501 xanthomonas
109	2.5	254	16	Q8G4Y2	Q8g4y2 bifidobacte	182	7	2.5	380	13	Q42317	Q42317 cyprinus ca
110	2.5	256	16	Q88RL7	Q88rl7 pseudomonas	183	7	2.5	380	16	Q9PGC2	Q9pgc2 xylella fas
111	2.5	260	16	Q9HT68	Q9ht68 pseudomonas	184	7	2.5	386	17	P95972	P95972 sulfolobus
112	2.5	260	16	Q87Q12	Q87q12 vibrio para	185	7	2.5	397	16	Q9AA49	Q9aa49 caulobacter
113	2.5	261	16	Q9PF05	Q9pf05 xylella fas	186	7	2.5	400	16	Q9KN80	Q9kn80 vibrio chol
114	2.5	261	16	Q8D8T3	Q8d8t3 vibrio vuln	187	7	2.5	401	2	Q51696	Q51696 brevundimon
115	2.5	261	16	Q87AL7	Q87al7 xylella fas	188	7	2.5	402	16	Q8A227	Q8a227 bacteroides
116	2.5	262	16	Q9KRF8	Q9krf8 vibrio chol	189	7	2.5	404	16	Q9WX27	Q9wx27 streptomyce
117	2.5	263	16	Q9KWS9	Q9kws9 vibrio algi	190	7	2.5	406	16	Q8ZMW4	Q8zwm4 salmonella
118	2.5	264	16	Q7VWV2	Q7vmv2 haemophilus	191	7	2.5	406	16	Q8ZAI6	Q8zai6 salmonella
119	2.5	265	16	Q83VH9	Q83vh9 treponema h	192	7	2.5	410	2	Q8G1S7	Q8g1s7 synechococc
120	2.5	265	16	Q9CWB7	Q9cmb7 pasteurella	193	7	2.5	411	10	Q8RVC7	Q8rvc7 chlamydomon
121	2.5	266	16	Q8Z7C4	Q8z7c4 salmonella	194	7	2.5	420	16	Q8XY93	Q8xy93 ralestonia s
122	2.5	268	16	Q8YAR0	Q8yaro escherichia	195	7	2.5	428	16	Q8XU06	Q8xu06 mycobacteri
123	2.5	268	16	Q83T26	Q83t26 salmonella	196	7	2.5	432	16	Q9RT67	Q9rt67 mycobacteri
124	2.5	269	16	Q86191	Q86191 erwania chr	197	7	2.5	434	16	Q8RA18	Q8ra18 chlamydomon
125	2.5	270	16	Q8DFC1	Q8dfc1 vibrio vuln	198	7	2.5	445	10	Q94IP6	Q94ip6 chlamydomon
126	2.5	270	16	Q8G8H9	Q8g8h9 erwania pyr	199	7	2.5	450	8	Q9AS89	Q9as89 anguilla an
127	2.5	270	16	Q81VM1	Q81vm1 bacillus an	200	7	2.5	460	16	Q8NSI8	Q8nsi8 monopterus
128	2.5	271	16	Q81I25	Q81i25 bacillus ce	201	7	2.5	468	5	Q19775	Q19775 caenorhabdi
129	2.5	271	16	Q26084	Q26084 helicobacte	202	7	2.5	478	11	Q8BLD9	Q8bl9 mus musculu
130	2.5	271	16	Q9ZJ45	Q9zj45 helicobacte	203	7	2.5	484	16	Q98HCO	Q98hc0 rhizobium 1
131	2.5	271	16	Q8Z527	Q8z527 yerisina pe	204	7	2.5	496	16	Q8CMY1	Q8cm1 staphylococ
132	2.5	273	16	Q92E25	Q92e25 listeria in	205	7	2.5	450	8	Q9AS89	Q9as89 ralestonia s
133	2.5	274	2	Q8KUB0	Q8kub0 vibrio fisc	206	7	2.5	460	16	Q8NSI8	Q8nsi8 corynebacte
134	2.5	274	2	Q7WVX6	Q7wvx6 borrelia ga	207	7	2.5	468	5	Q19775	Q19775 caenorhabdi
135	2.5	274	16	Q8ZG10	Q8zgi0 yerisina pe	208	7	2.5	478	4	Q8NEQ8	Q8neq8 homo sapien
136	2.5	275	16	Q8ZBF9	Q8zbf9 streptomyce	209	7	2.5	484	11	Q8BLD9	Q8bl9 mus musculu
137	2.5	276	16	Q9Z8M0	Q9z8m0 listeria in	210	7	2.5	496	16	Q8CMY1	Q8cm1 staphylococ
138	2.5	276	16	Q8Y4M0	Q8y4m0 listeria mo	211	7	2.5	527	5	Q9XZ04	Q9xz04 drosophila
139	2.5	279	16	Q8ENU4	Q8enu4 oceanobacili	212	7	2.5	527	5	Q9VC35	Q9vc35 drosophila
140	2.5	282	16	Q9BMD2	Q9bmd2 plasmodium	213	7	2.5	527	5	Q9VC35	Q9vc35 drosophila
141	2.5	282	16	Q89HG5	Q89hg5 bradyrhizob	214	7	2.5	527	5	Q86PC4	Q86pc4 drosophila
142	2.5	284	2	Q93EAS	Q93eas rhizobium 1	215	7	2.5	527	5	Q9XZ04	Q9xz04 drosophila
143	2.5	284	16	Q98DA1	Q98da1 rhizobium 1	216	7	2.5	527	5	Q9VC35	Q9vc35 drosophila
144	2.5	286	16	Q9CIN8	Q9cin8 lactococcus	217	7	2.5	530	17	Q8PYV7	Q8pyv7 methanosarc
145	2.5	286	16	Q9CIN8	Q9cin8 lactococcus	218	7	2.5	530	17	Q8PYV7	Q8pyv7 methanosarc
146	2.5	287	2	Q9JPI7	Q9jpi7 neisseria m	219	7	2.5	541	10	Q84TG1	Q84tg1 arabidopsis
147	2.5	287	2	Q9JPI7	Q9jpi7 neisseria m	220	7	2.5	541	10	Q84TG1	Q84tg1 arabidopsis
148	2.5	289	16	Q8ZMS5	Q8zms5 streptomyce	221	7	2.5	545	10	Q8G8T1	Q8g8t1 arabidopsis
149	2.5	290	13	Q800B7	Q800b7 rana rugosa	222	7	2.5	549	16	Q97M53	Q97m53 clostridium
150	2.5	290	13	Q800B6	Q800b6 rana rugosa	223	7	2.5	554	5	Q93174	Q93174 caenorhabdi
151	2.5	290	13	Q7ZT24	Q7zt24 rana rugosa	224	7	2.5	557	10	Q8VZB2	Q8vzb2 arabidopsis
152	2.5	291	3	Q9C445	Q9c445 penicillium	225	7	2.5	563	16	Q891J9	Q891j9 clostridium
153	2.5	295	9	Q7YSW6	Q7ysw6 bacteriophag	226	7	2.5	566	5	Q810H9	Q810h9 drosophila
154	2.5	296	16	Q7UOG5	Q7uog5 rhodospirill	227	7	2.5	572	16	Q9S1S8	Q9s1s8 streptomyce
155	2.5	298	13	Q8SCH5	Q8sch5 bos taurus	228	7	2.5	580	16	Q5S856	Q5s856 synechocyst
156	2.5	298	13	Q9YTC4	Q9ytc4 rana rugosa	229	7	2.5	581	16	Q89NWI	Q89nwi bradyrhizob
157	2.5	298	13	Q9ZRH2	Q9zrh2 rana rugosa	230	7	2.5	585	16	Q8A7X1	Q8a7x1 bacteroides
158	2.5	298	13	Q9PRH1	Q9prh1 rana rugosa	231	7	2.5	586	16	Q97TJ5	Q97tj5 clostridium
159	2.5	298	13	Q8AYW3	Q8ayw3 gallus gall	232	7	2.5	598	5	Q8PII0	Q8pii0 drosophila
160	2.5	298	13	Q919V9	Q919v9 xenopus lae	233	7	2.5	599	5	Q25310	Q25310 leishmania
161	2.5	298	13	Q8JHI0	Q8jhi0 brachydanio	234	7	2.5	600	5	Q01530	Q01530 trypanosoma
162	2.5	300	5	Q01813	Q01813 caenorhabdi	235	7	2.5	601	16	Q8AB55	Q8ab55 bacteroides

236	7	2.5	610	5	Q19848	Q19848 caenorhabdi	309	7	2.5	18412	13	Q7Z261	Q7Z261 brachydanio
237	7	2.5	610	10	Q9LYN3	Q9LYN3 arabidopsis	310	6	2.2	20	2	Q9R4K6	Q9R4K6 staphylococ
238	7	2.5	611	17	Q9UZU0	Q9UZU0 pyrococcus	311	6	2.2	20	15	Q9DYR0	Q9DYR0 human immun
239	7	2.5	616	16	Q9KQK4	Q9KQK4 vibrio chol	312	6	2.2	33	4	Q16148	Q16148 homo sapien
240	7	2.5	634	5	Q8STZ1	Q8STZ1 encephalito	313	6	2.2	35	16	Q8CQ13	Q8CQ13 staphylococ
241	7	2.5	634	5	Q8ST44	Q8ST44 encephalito	314	6	2.2	42	10	P93072	P93072 bambusa sp.
242	7	2.5	638	16	Q8RT72	Q8RT72 fusobacteri	315	6	2.2	44	4	Q13013	Q13013 homo sapien
243	7	2.5	652	16	Q97OT2	Q97OT2 streptococ	316	6	2.2	44	16	Q8CSY2	Q8CSY2 staphylococ
244	7	2.5	652	16	Q8DP59	Q8DP59 streptococ	317	6	2.2	44	16	Q7W6T7	Q7W6T7 bordetella
245	7	2.5	653	10	Q9LSA0	Q9LSA0 arabidopsis	318	6	2.2	45	1	Q7W6U1	Q7W6U1 bordetella
246	7	2.5	660	16	Q8Y140	Q8Y140 ralestonia s	319	6	2.2	45	2	P83309	P83309 comanonas a
247	7	2.5	668	5	Q3V8D2	Q3V8D2 dirosophila	320	6	2.2	45	4	Q12847	Q12847 homo sapien
248	7	2.5	668	5	Q8SZH5	Q8SZH5 dirosophila	321	6	2.2	45	4	Q12847	Q12847 homo sapien
249	7	2.5	670	16	Q5V138	Q5V138 dirosophila	322	6	2.2	48	2	Q9X9H7	Q9X9H7 yersinia en
250	7	2.5	691	3	Q3URQ2	Q3URQ2 saccharomyc	323	6	2.2	50	16	Q8X3B6	Q8X3B6 escherichia
251	7	2.5	691	16	Q34523	Q34523 helicobacte	324	6	2.2	50	16	Q8Z8R0	Q8Z8R0 streptomyc
252	7	2.5	692	16	Q8CY25	Q8CY25 corynebacte	325	6	2.2	56	10	Q9LJ00	Q9LJ00 oryza sativ
253	7	2.5	694	16	Q9X120	Q9X120 thermocoga	326	6	2.2	57	9	Q8SDJ4	Q8SDJ4 streptococ
254	7	2.5	696	16	Q9ZVD4	Q9ZVD4 rhizobium m	327	6	2.2	57	10	Q7X1N8	Q7X1N8 oryza sativ
255	7	2.5	703	16	Q7UXT8	Q7UXT8 rhodospirell	328	6	2.2	60	2	Q9K1Z7	Q9K1Z7 edwardsiell
256	7	2.5	718	4	Q9HCD1	Q9HCD1 homo sapien	329	6	2.2	60	3	Q17711	Q17711 trichoderma
257	7	2.5	721	5	Q9VMU5	Q9VMU5 dirosophila	330	6	2.2	62	16	Q89VU7	Q89VU7 bradyrhizob
258	7	2.5	721	10	Q941P7	Q941P7 chlamydomon	331	6	2.2	62	16	Q928D7	Q928D7 listeria in
259	7	2.5	723	16	Q8DA33	Q8DA33 vibrio vuln	332	6	2.2	64	16	Q9S1E1	Q9S1E1 yersinia pe
260	7	2.5	734	12	Q8OY51	Q8OY51 heterobasid	333	6	2.2	66	2	Q9AP90	Q9AP90 uncultured
261	7	2.5	741	16	Q8EIV5	Q8EIV5 shewanella	334	6	2.2	67	2	Q85232	Q85232 staphylococ
262	7	2.5	749	5	Q9W4E0	Q9W4E0 dirosophila	335	6	2.2	67	5	Q20433	Q20433 caenorhabdi
263	7	2.5	762	17	Q976U7	Q976U7 sulfolobus	336	6	2.2	67	10	Q852U3	Q852U3 brassica ju
264	7	2.5	776	2	Q8RN57	Q8RN57 campylobact	337	6	2.2	68	16	Q83E18	Q83E18 coxiella bu
265	7	2.5	796	4	Q8NC87	Q8NC87 homo sapien	338	6	2.2	67	16	Q9RRN6	Q9RRN6 deinococcus
266	7	2.5	796	5	Q86NX5	Q86NX5 dirosophila	339	6	2.2	68	16	Q9KBB4	Q9KBB4 bacillus ha
267	7	2.5	802	2	Q8RN16	Q8RN16 campylobact	340	6	2.2	68	16	Q8FRCS	Q8FRCS corynebacte
268	7	2.5	812	16	Q8FWG2	Q8FWG2 corynebacte	341	6	2.2	69	2	Q52022	Q52022 staphylococ
269	7	2.5	860	16	Q8NM83	Q8NM83 corynebacte	342	6	2.2	70	10	Q7XU07	Q7XU07 oryza sativ
270	7	2.5	867	2	Q7WZ88	Q7WZ88 nonomureta	343	6	2.2	70	16	Q8ZDW4	Q8ZDW4 yersinia pe
271	7	2.5	867	16	Q8Y6W6	Q8Y6W6 listeria mo	344	6	2.2	71	7	Q9GJ66	Q9GJ66 salmo trutt
272	7	2.5	875	16	Q8UCC1	Q8UCC1 agrobacteri	345	6	2.2	71	7	Q9GJ63	Q9GJ63 salmo trutt
273	7	2.5	905	16	Q894I2	Q894I2 clostridium	346	6	2.2	71	7	Q9GJ65	Q9GJ65 salmo trutt
274	7	2.5	913	16	Q83GL1	Q83GL1 tropheryma	347	6	2.2	71	7	Q9GJ64	Q9GJ64 salmo trutt
275	7	2.5	938	10	Q9AM83	Q9AM83 guillardia	348	6	2.2	71	16	Q87NS5	Q87NS5 vibrio para
276	7	2.5	938	10	Q83HL6	Q83HL6 tropheryma	349	6	2.2	72	2	Q52023	Q52023 staphylococ
277	7	2.5	965	5	Q8WVZ6	Q8WVZ6 cryptospori	350	6	2.2	72	12	Q8Q0Q9	Q8Q0Q9 camelipox vi
278	7	2.5	967	16	Q98GNS	Q98GNS rhizobium 1	351	6	2.2	73	16	Q68801	Q68801 pseudomonas
279	7	2.5	1012	10	Q7XLU5	Q7XLU5 oryza sativ	352	6	2.2	73	16	Q881I6	Q881I6 pseudomonas
280	7	2.5	1036	16	Q9KVI2	Q9KVI2 vibrio chol	353	6	2.2	74	17	Q8ZUW3	Q8ZUW3 pyrobaculum
281	7	2.5	1039	16	Q8DDC6	Q8DDC6 vibrio vuln	354	6	2.2	76	9	Q8H9X8	Q8H9X8 pseudomonas
282	7	2.5	1040	16	Q87TNI	Q87TNI arabidopsis	355	6	2.2	76	10	Q9LFR0	Q9LFR0 arabidopsis
283	7	2.5	1095	10	Q9CTH5	Q9CTH5 arabidopsis	356	6	2.2	79	12	Q8Q078	Q8Q078 camelipox vi
284	7	2.5	1118	16	Q97E71	Q97E71 clostridium	357	6	2.2	81	2	Q8S411	Q8S411 streptococ
285	7	2.5	1130	5	Q9XTI9	Q9XTI9 caenorhabdi	358	6	2.2	82	16	Q8XDP5	Q8XDP5 escherichia
286	7	2.5	1130	16	Q7UT96	Q7UT96 rhodospirell	359	6	2.2	82	17	Q8TRV0	Q8TRV0 methanopyru
287	7	2.5	1130	16	Q8DE74	Q8DE74 vibrio vuln	360	6	2.2	83	6	Q864U5	Q864U5 bos taurus
288	7	2.5	1171	16	Q8NSED	Q8NSED corynebacte	361	6	2.2	83	16	Q8XP54	Q8XP54 clostridium
289	7	2.5	1233	16	Q8NSED	Q8NSED corynebacte	362	6	2.2	85	12	Q8V6P2	Q8V6P2 halovirus h
290	7	2.5	1240	4	Q8NFM2	Q8NFM2 homo sapien	363	6	2.2	85	12	Q7TDL1	Q7TDL1 halovirus h
291	7	2.5	1283	10	Q8LIT4	Q8LIT4 oryza sativ	364	6	2.2	85	16	Q9JZX3	Q9JZX3 neisseria m
292	7	2.5	1333	5	Q24262	Q24262 dirosophila	365	6	2.2	86	3	Q9US10	Q9US10 schizosacch
293	7	2.5	1431	10	Q7XVY4	Q7XVY4 oryza sativ	366	6	2.2	86	16	Q8DIM5	Q8DIM5 synecococ
294	7	2.5	1496	2	Q93IE8	Q93IE8 actinobacil	367	6	2.2	87	10	Q8L9S3	Q8L9S3 arabidopsis
295	7	2.5	1498	2	Q91448	Q91448 arthropacte	368	6	2.2	88	17	Q8ZVY4	Q8ZVY4 pyrobaculum
296	7	2.5	1615	4	Q9NYS8	Q9NYS8 homo sapien	369	6	2.2	90	16	Q8XSV2	Q8XSV2 escherichia
297	7	2.5	1616	4	Q8NEV4	Q8NEV4 homo sapien	370	6	2.2	90	16	Q8XSV2	Q8XSV2 escherichia
298	7	2.5	1784	5	Q9VES4	Q9VES4 dirosophila	371	6	2.2	90	16	Q8FB46	Q8FB46 mycobacteri
299	7	2.5	1795	16	Q9ZUF7	Q9ZUF7 rickettsia	372	6	2.2	90	16	Q8XGB5	Q8XGB5 salmonella
300	7	2.5	1919	10	Q8LRK9	Q8LRK9 arabidopsis	373	6	2.2	90	16	Q7TWM0	Q7TWM0 mycobacteri
301	7	2.5	1994	10	Q9LPI9	Q9LPI9 pseudomonas	374	6	2.2	91	16	Q83B37	Q83B37 coxiella bu
302	7	2.5	2147	2	Q9L350	Q9L350 pseudomonas	375	6	2.2	92	2	Q63391	Q63391 methylomona
303	7	2.5	2154	16	Q9HVG6	Q9HVG6 pseudomonas	376	6	2.2	92	5	Q18742	Q18742 caenorhabdi
304	7	2.5	4066	2	Q8XLL5	Q8XLL5 streptomyc	377	6	2.2	92	11	Q8BLI4	Q8BLI4 mus musculu
305	7	2.5	4360	2	Q30764	Q30764 streptomyc	378	6	2.2	93	16	Q7VIN3	Q7VIN3 helicobacte
306	7	2.5	4688	16	Q9FQ08	Q9FQ08 ureaplasma	379	6	2.2	93	17	Q26421	Q26421 methanobact
307	7	2.5	6310	16	Q88FP2	Q88FP2 pseudomonas	380	6	2.2	95	11	Q8K2X1	Q8K2X1 mus musculu
308	7	2.5	6396	2	Q9KID7	Q9KID7 streptomyc	381	6	2.2	95	16	Q83JC6	Q83JC6 shigella fl

382	6	2.2	95	16	082WV6	082WV6 nitrosomona	455	6	2.2	126	16	085D50	085D50 vibrio vuln
383	6	2.2	96	9	Q37904	Q37904 enterobacte	456	6	2.2	127	5	081P12	081P12 drosophila
384	6	2.2	96	11	063156	063156 rattus norv	457	6	2.2	127	10	09EXR3	09EXR3 nicotiana t
385	6	2.2	96	16	07V441	07V441 prochloroco	458	6	2.2	127	10	09EXR2	09EXR2 nicotiana t
386	6	2.2	97	6	086214	086214 bos taurus	459	6	2.2	127	12	091G17	091G17 epiphyas po
387	6	2.2	97	16	098U50	098U50 rhizobium 1	460	6	2.2	127	16	082GHS	082GHS streptomyc
388	6	2.2	98	5	001649	001649 drosophila	461	6	2.2	128	2	084B92	084B92 gamma-prote
389	6	2.2	98	6	095N84	095N84 ovis aries	462	6	2.2	128	2	084BL3	084BL3 gamma-prote
391	6	2.2	99	16	Q9XB24	Q9XB24 bacterioph	463	6	2.2	128	3	094597	094597 schistosach
392	6	2.2	99	16	09YFC2	09YFC2 bacillus ha	464	6	2.2	128	12	089084	089084 variola vir
393	6	2.2	99	16	082363	082363 brucella me	465	6	2.2	128	16	097TK4	097TK4 clostridium
394	6	2.2	101	16	08RJK0	08RJK0 nitrosococ	466	6	2.2	128	16	097CY3	097CY3 clostridium
395	6	2.2	101	16	08DSC9	08DSC9 streptococ	467	6	2.2	128	16	08U9C9	08U9C9 agrobacteri
396	6	2.2	101	16	Q7V314	Q7V314 prochloroco	468	6	2.2	128	16	086GY2	086GY2 pseudomonas
397	6	2.2	102	10	Q8LEU7	Q8LEU7 arabidopsis	469	6	2.2	129	11	063911	063911 mus sp. tap
398	6	2.2	102	10	Q8W453	Q8W453 arabidopsis	470	6	2.2	129	17	08U116	08U116 pyrococcus
399	6	2.2	102	16	Q8DM61	Q8DM61 streptococ	471	6	2.2	130	2	085202	085202 streptococ
400	6	2.2	102	16	Q9YV71	Q9YV71 streptococ	472	6	2.2	130	2	08RSP3	08RSP3 streptococ
401	6	2.2	103	16	Q7WUB7	Q7WUB7 escherichia	473	6	2.2	130	2	09F518	09F518 escherichia
402	6	2.2	103	16	Q7VOC9	Q7VOC9 candidatus	474	6	2.2	130	2	09S347	09S347 prevotella
403	6	2.2	104	16	092DX8	092DX8 listeria in	475	6	2.2	130	16	08D8A5	08D8A5 vibrio vuln
404	6	2.2	104	16	08Y959	08Y959 listeria mo	476	6	2.2	130	16	087UY9	087UY9 pseudomonas
405	6	2.2	105	6	Q93R83	Q93R83 sus scrofa	477	6	2.2	130	16	087N42	087N42 vibrio para
406	6	2.2	105	6	Q9T5Y3	Q9T5Y3 sus scrofa	478	6	2.2	132	6	0862E8	0862E8 bos taurus
407	6	2.2	105	16	Q7WE14	Q7WE14 bordetella	479	6	2.2	132	16	09C7Q3	09C7Q3 pasteurella
408	6	2.2	105	16	Q7M566	Q7M566 bordetella	480	6	2.2	132	16	08FWR3	08FWR3 brucella su
409	6	2.2	105	16	Q7VSO1	Q7VSO1 bordetella	481	6	2.2	132	16	0821G2	0821G2 chlamydomo
410	6	2.2	106	10	P93580	P93580 stipa virid	482	6	2.2	133	2	07WWE6	07WWE6 yersinia al
411	6	2.2	106	10	Q9M1Y4	Q9M1Y4 arabidopsis	483	6	2.2	133	5	Q21887	Q21887 caenorhabd
412	6	2.2	107	16	Q88VP4	Q88VP4 lacobacilli	484	6	2.2	133	11	Q921B5	Q921B5 mus musculu
413	6	2.2	108	16	Q8XK16	Q8XK16 lacobacilli	485	6	2.2	133	16	08FMP5	08FMP5 corynebacte
414	6	2.2	108	16	Q8EM18	Q8EM18 oceanobacill	486	6	2.2	133	16	08DCH9	08DCH9 vibrio vuln
415	6	2.2	108	16	Q8D9H7	Q8D9H7 vibrio vuln	487	6	2.2	134	5	015615	015615 entamoeba h
416	6	2.2	109	2	Q9EUV9	Q9EUV9 listeria in	488	6	2.2	134	16	087TK9	087TK9 vibrio para
417	6	2.2	109	16	Q8YBZ2	Q8YBZ2 brucella me	489	6	2.2	135	2	09F7U0	09F7U0 helicobacte
418	6	2.2	110	5	Q9YUO0	Q9YUO0 drosophila	490	6	2.2	135	5	09G8Q7	09G8Q7 habesita big
419	6	2.2	110	16	Q9K6V9	Q9K6V9 bacillus ha	491	6	2.2	135	8	Q7Y2N1	Q7Y2N1 celatoblat
420	6	2.2	110	16	Q9F5L4	Q9F5L4 brachyriroz	492	6	2.2	135	8	Q7Y1A0	Q7Y1A0 celatoblat
421	6	2.2	111	10	080578	080578 arabidopsis	493	6	2.2	135	8	Q7Y196	Q7Y196 celatoblat
422	6	2.2	111	16	Q8NU16	Q8NU16 corynebacte	494	6	2.2	135	8	Q7Y195	Q7Y195 celatoblat
423	6	2.2	112	16	P96695	P96695 bacillus su	495	6	2.2	135	8	Q7Y193	Q7Y193 celatoblat
424	6	2.2	112	16	Q8FST0	Q8FST0 corynebacte	496	6	2.2	135	8	Q7Y192	Q7Y192 celatoblat
425	6	2.2	112	16	Q7USG0	Q7USG0 rhodospirill	497	6	2.2	135	8	Q7Y191	Q7Y191 celatoblat
426	6	2.2	113	3	060048	060048 pleurotus o	498	6	2.2	135	8	Q7Y190	Q7Y190 celatoblat
427	6	2.2	113	11	Q9D4T7	Q9D4T7 mus musculu	499	6	2.2	135	8	Q7Y189	Q7Y189 celatoblat
428	6	2.2	113	17	Q8YB5	Q8YB5 pyrobaculum	500	6	2.2	135	8	Q7Y188	Q7Y188 celatoblat
429	6	2.2	114	9	Q856B1	Q856B1 mycobacteri	501	6	2.2	135	8	Q7Y186	Q7Y186 celatoblat
430	6	2.2	114	11	Q8CAK0	Q8CAK0 mus musculu	502	6	2.2	135	8	Q7Y185	Q7Y185 celatoblat
431	6	2.2	114	11	Q9NM00	Q9NM00 mus musculu	503	6	2.2	135	8	Q7Y183	Q7Y183 celatoblat
432	6	2.2	115	12	Q9EMW7	Q9EMW7 amaestra moo	504	6	2.2	135	8	Q7Y182	Q7Y182 celatoblat
433	6	2.2	115	5	Q9VAK8	Q9VAK8 drosophila	505	6	2.2	135	8	Q7Y181	Q7Y181 platyoster
434	6	2.2	115	16	Q8EDV4	Q8EDV4 shewanella	506	6	2.2	135	8	Q7Y180	Q7Y180 drymaplanet
435	6	2.2	115	16	Q7WNV4	Q7WNV4 bordetella	507	6	2.2	135	8	Q7Y179	Q7Y179 celatoblat
436	6	2.2	115	16	Q7VUV5	Q7VUV5 bordetella	508	6	2.2	135	8	Q7Y178	Q7Y178 celatoblat
437	6	2.2	115	17	Q9V100	Q9V100 pyrococcus	509	6	2.2	135	16	Q9K1N2	Q9K1N2 neisseria m
438	6	2.2	116	10	Q42353	Q42353 arabidopsis	510	6	2.2	135	16	Q93JH3	Q93JH3 streptomyc
439	6	2.2	116	16	Q91786	Q91786 pseudomonas	511	6	2.2	135	16	Q7WLC6	Q7WLC6 bordetella
440	6	2.2	117	16	Q83HY2	Q83HY2 coxiella hu	512	6	2.2	136	5	Q9NH93	Q9NH93 bombyx mori
441	6	2.2	117	2	Q84F43	Q84F43 clostridium	513	6	2.2	136	5	062434	062434 caenorhabd
442	6	2.2	120	2	Q92N12	Q92N12 staphylococ	514	6	2.2	136	10	Q9N104	Q9N104 ovis aries
443	6	2.2	120	16	Q87CV5	Q87CV5 xylella fas	515	6	2.2	137	6	Q940S8	Q940S8 rosa hybrid
444	6	2.2	120	16	Q7UPT2	Q7UPT2 rhodospirill	516	6	2.2	137	16	Q94WU1	Q94WU1 periplaneta
445	6	2.2	121	16	Q9JU23	Q9JU23 neisseria m	517	6	2.2	137	16	Q8EYV6	Q8EYV6 shewanella
446	6	2.2	121	16	Q81UV8	Q81UV8 bacillus an	518	6	2.2	137	16	Q87J29	Q87J29 vibrio para
447	6	2.2	122	5	Q8ST62	Q8ST62 encephalito	519	6	2.2	137	16	Q831Z3	Q831Z3 shigella fl
448	6	2.2	122	5	Q8SC07	Q8SC07 streptococ	520	6	2.2	138	16	Q95Y93	Q95Y93 caenorhabd
449	6	2.2	122	9	Q21992	Q21992 caenorhabd	521	6	2.2	138	16	Q95Y93	Q95Y93 caenorhabd
450	6	2.2	124	16	Q9W353	Q9W353 staphylococ	522	6	2.2	139	10	Q7X7F8	Q7X7F8 oryza sativ
451	6	2.2	125	2	087959	087959 salmonella	523	6	2.2	139	10	Q89100	Q89100 bradyrhizob
452	6	2.2	126	16	08YPY1	08YPY1 anabaena sp	524	6	2.2	139	16		
453	6	2.2					525	6	2.2				
454	6	2.2					526	6	2.2				
							527	6	2.2				

528	6	2.2	139	16	Q7V985	Q7V985 prochloroc	601	6	2.2	157	2	Q9KRD5	Q9KRD5 streptomyc
529	6	2.2	140	8	Q8WAC3	Q8WAC3 ligia yitie	602	6	2.2	157	1	Q9J904	Q9J904 caenorhabdi
530	6	2.2	141	16	Q8PIA4	Q8PIA4 streptococ	603	6	2.2	157	11	Q8QVB5	Q8QVB5 mus musculu
531	6	2.2	141	16	Q8K788	Q8K788 streptococ	604	6	2.2	157	16	Q9BB79	Q9BB79 rhizobium i
532	6	2.2	141	16	Q8BYE1	Q8BYE1 lactobacill	605	6	2.2	157	16	Q92S89	Q92S89 rhizobium m
533	6	2.2	142	2	Q9ZAE8	Q9ZAE8 streptomyc	606	6	2.2	157	16	Q89DR1	Q89DR1 bradyrhizob
534	6	2.2	142	8	Q9T250	Q9T250 phycophthor	607	6	2.2	157	16	Q82L06	Q82L06 streptomyc
535	6	2.2	142	16	Q8XL69	Q8XL69 clostridium	608	6	2.2	158	2	Q84EB9	Q84EB9 gamma-prote
536	6	2.2	142	16	Q7UJ87	Q7UJ87 rhodopirell	609	6	2.2	158	4	Q8IX95	Q8IX95 homo sapien
537	6	2.2	143	9	Q9XJ00	Q9XJ00 streptococ	610	6	2.2	158	4	Q8WXS7	Q8WXS7 diosiphila
538	6	2.2	143	16	Q9Z2C1	Q9Z2C1 rhizobium m	611	6	2.2	158	13	Q9PVJ3	Q9PVJ3 typhlops-xi
539	6	2.2	143	16	Q8BDJ2	Q8BDJ2 pseudomonas	612	6	2.2	158	16	Q9FM77	Q9FM77 xylella fas
540	6	2.2	143	17	Q8PVI4	Q8PVI4 methanosarc	613	6	2.2	158	16	Q9I1B5	Q9I1B5 pseudomonas
541	6	2.2	144	2	Q9LAI2	Q9LAI2 pasteurella	614	6	2.2	158	16	Q8XDC7	Q8XDC7 escherichia
542	6	2.2	144	16	Q9CPF9	Q9CPF9 pasteurella	615	6	2.2	158	16	Q8NLD9	Q8NLD9 corynebacte
543	6	2.2	144	16	Q8RRM0	Q8RRM0 fusobacteri	616	6	2.2	158	16	Q8AB59	Q8AB59 bacteroides
544	6	2.2	144	16	Q8RGE8	Q8RGE8 fusobacteri	617	6	2.2	158	16	Q8TBD0	Q8TBD0 xylella fas
545	6	2.2	144	16	Q8BED4	Q8BED4 pseudomonas	618	6	2.2	159	2	Q9X2T7	Q9X2T7 macrococcus
546	6	2.2	144	16	Q8XGJ6	Q8XGJ6 salmoneila	619	6	2.2	159	9	Q858Y7	Q858Y7 bacterioph
547	6	2.2	144	16	Q83R76	Q83R76 shigella fl	620	6	2.2	159	12	P87547	P87547 babanki vir
548	6	2.2	144	17	Q58I42	Q58I42 pyrococcus	621	6	2.2	159	16	Q8R7B9	Q8R7B9 listeria mo
549	6	2.2	145	16	Q8PR60	Q8PR60 xanthomonas	622	6	2.2	159	16	Q8RH55	Q8RH55 fusobacteri
550	6	2.2	145	16	Q8PW79	Q8PW79 xanthomonas	623	6	2.2	159	16	Q8ZCJ5	Q8ZCJ5 yersinia pe
551	6	2.2	146	6	Q8MJ96	Q8MJ96 equus cabal	624	6	2.2	159	16	Q8R1N7	Q8R1N7 bacillus an
552	6	2.2	146	5	Q9VDS7	Q9VDS7 diosiphila	625	6	2.2	159	16	Q7UPB1	Q7UPB1 rhodopirell
553	6	2.2	147	5	Q26505	Q26505 scapharca i	626	6	2.2	160	5	Q8WS63	Q8WS63 glosina mo
554	6	2.2	147	6	Q28523	Q28523 macaca mula	627	6	2.2	160	11	Q8C7Y8	Q8C7Y8 mus musculu
555	6	2.2	147	16	Q8ZNU2	Q8ZNU2 salmoneila	628	6	2.2	160	16	Q8PJ32	Q8PJ32 xanthomonas
556	6	2.2	147	16	Q89P88	Q89P88 bradyrhizob	629	6	2.2	160	16	Q8R8Y4	Q8R8Y4 leptospira
557	6	2.2	147	16	Q7WYX6	Q7WYX6 bordetella	630	6	2.2	161	2	Q8S011	Q8S011 bordetella bu
558	6	2.2	147	16	Q7VX30	Q7VX30 bordetella	631	6	2.2	161	8	Q8HGV9	Q8HGV9 typhlatya m
559	6	2.2	147	16	Q7VBE6	Q7VBE6 prochloroc	632	6	2.2	161	8	Q8HGV8	Q8HGV8 typhlatya m
560	6	2.2	148	1	Q93698	Q93698 sulfolobac	633	6	2.2	161	8	Q8HGV7	Q8HGV7 typhlatya m
561	6	2.2	148	2	Q9RB19	Q9RB19 acinetobac	634	6	2.2	161	8	Q8HFT4	Q8HFT4 typhlatya p
562	6	2.2	148	16	Q9A2M3	Q9A2M3 caulopacter	635	6	2.2	161	8	Q8HFT3	Q8HFT3 typhlatya p
563	6	2.2	148	16	Q97ORA	Q97ORA streptococ	636	6	2.2	161	8	Q8HFT2	Q8HFT2 typhlatya p
564	6	2.2	148	16	Q8BDP1	Q8BDP1 pseudomonas	637	6	2.2	161	8	Q8HBA5	Q8HBA5 typhlatya s
565	6	2.2	148	16	Q87YK6	Q87YK6 pseudomonas	638	6	2.2	161	8	Q8HBS5	Q8HBS5 typhlatya m
566	6	2.2	148	16	Q87YK6	Q87YK6 xylella fas	639	6	2.2	161	8	Q8HBS4	Q8HBS4 typhlatya m
567	6	2.2	149	2	Q84HX2	Q84HX2 micromonos	640	6	2.2	161	16	Q86G55	Q86G55 pseudomonas
568	6	2.2	149	2	Q7WME7	Q7WME7 yersinia al	641	6	2.2	161	16	Q83JH6	Q83JH6 enterococcu
569	6	2.2	149	15	Q7SSW7	Q7SSW7 chimpanzee	642	6	2.2	162	2	Q9PDS9	Q9PDS9 uncultured
570	6	2.2	149	16	Q8XCV4	Q8XCV4 bruceella me	643	6	2.2	162	12	P87710	P87710 rotavirus a
571	6	2.2	149	16	Q8XC10	Q8XC10 escherichia	644	6	2.2	162	12	Q02476	Q02476 poa semilat
572	6	2.2	149	16	Q8XUL8	Q8XUL8 ralsstonia s	645	6	2.2	162	16	Q8PGC5	Q8PGC5 xanthomonas
573	6	2.2	149	16	Q8XUL8	Q8XUL8 ralsstonia s	646	6	2.2	163	2	Q93IH5	Q93IH5 wolstonia s
574	6	2.2	149	16	Q8PH48	Q8PH48 xanthomonas	647	6	2.2	163	5	Q17155	Q17155 babactia.11
575	6	2.2	149	16	Q8FOT2	Q8FOT2 leptospira	648	6	2.2	163	16	Q7V0X2	Q7V0X2 prochloroc
576	6	2.2	149	17	Q8PX90	Q8PX90 methanosarc	649	6	2.2	164	12	Q9EMV7	Q9EMV7 amascta moo
577	6	2.2	150	2	Q84EAO	Q84EAO gamma-proce	650	6	2.2	164	17	Q8ZT11	Q8ZT11 pyrobaculum
578	6	2.2	150	16	Q981Z3	Q981Z3 rhizobium i	651	6	2.2	165	5	Q868X8	Q868X8 diosiphila
579	6	2.2	151	2	Q9AHX4	Q9AHX4 carsonella	652	6	2.2	165	16	Q7USL9	Q7USL9 rhodopirell
580	6	2.2	151	16	Q8ABX8	Q8ABX8 bacteroides	653	6	2.2	166	5	Q9VWV7	Q9VWV7 diosiphila
581	6	2.2	151	10	Q89NU0	Q89NU0 bradyrhizob	654	6	2.2	166	16	Q8R7A3	Q8R7A3 thermoaer
582	6	2.2	152	16	Q7XSU0	Q7XSU0 oryza sativ	655	6	2.2	166	16	Q8KCT4	Q8KCT4 chlorobiu
583	6	2.2	152	16	Q97Q98	Q97Q98 streptococ	656	6	2.2	167	4	Q9P0R0	Q9P0R0 homo sapien
584	6	2.2	152	16	Q8NYP4	Q8NYP4 streptococ	657	6	2.2	167	4	Q15183	Q15183 homo sapien
585	6	2.2	153	16	Q92S17	Q92S17 rhizobium m	658	6	2.2	167	5	Q77278	Q77278 diosiphila
586	6	2.2	153	16	Q9ADV5	Q9ADV5 streptomyc	659	6	2.2	167	10	Q84K48	Q84K48 cucumis mel
587	6	2.2	153	16	Q92VL2	Q92VL2 rhizobium m	660	6	2.2	167	11	Q8BHG9	Q8BHG9 mus musculu
588	6	2.2	154	5	Q86MP4	Q86MP4 caenorhabdi	661	6	2.2	167	16	Q7WNO4	Q7WNO4 bordetella
589	6	2.2	154	16	Q82MZ7	Q82MZ7 streptomyc	662	6	2.2	167	16	Q7WBI2	Q7WBI2 bordetella
590	6	2.2	155	5	Q7UZR3	Q7UZR3 prochloroc	663	6	2.2	168	2	Q8GCG0	Q8GCG0 escherichia
591	6	2.2	155	16	Q21024	Q21024 caenorhabdi	664	6	2.2	168	4	Q9UFW8	Q9UFW8 homo sapien
592	6	2.2	155	16	Q98NS7	Q98NS7 rhizobium i	665	6	2.2	168	10	Q8H7F3	Q8H7F3 atapidopsi
593	6	2.2	156	3	Q9AH07	Q9AH07 rhodococcus	666	6	2.2	168	16	Q97SW0	Q97SW0 streptococ
594	6	2.2	156	3	Q05660	Q05660 saccharomyc	667	6	2.2	168	16	Q8X3P8	Q8X3P8 escherichia
595	6	2.2	156	13	Q9PVJ6	Q9PVJ6 eumeces ine	668	6	2.2	168	16	Q8YRT5	Q8YRT5 bruceella me
596	6	2.2	156	16	Q8XES7	Q8XES7 escherichia	669	6	2.2	168	16	Q8HFP3	Q8HFP3 escherichia
597	6	2.2	156	16	Q8RZ27	Q8RZ27 bruceella su	670	6	2.2	168	16	Q8CZ86	Q8CZ86 streptococ
598	6	2.2	156	16	Q8CWM5	Q8CWM5 escherichia	671	6	2.2	168	16	Q7UCM9	Q7UCM9 shigella fl
599	6	2.2	156	17	Q97U00	Q97U00 sulfolobus	672	6	2.2	168	16	Q9HNF5	Q9HNF5 halobacteri
600	6	2.2	157	2	Q847J0	Q847J0 pseudomonas	673	6	2.2	168	17		

674	6	2.2	168	17	Q8TUP3	Q8tup3 methanosarc	747	6	2.2	181	16	Q813P1	Q813p1 bacillus ce
675	6	2.2	169	2	Q9AP50	Q9ap50 streptococc	748	6	2.2	181	16	Q7MG64	Q7mg64 bordetella
676	6	2.2	169	2	Q9K3A8	Q9k3a8 enterococcu	749	6	2.2	181	16	Q7W4Y6	Q7w4y6 bordetella
677	6	2.2	169	2	Q84E97	Q84e97 gamma-prote	750	6	2.2	182	10	Q8S0E3	Q8s0e3 oryza sativ
678	6	2.2	169	2	Q7X0P0	Q7x0p0 uncultured	751	6	2.2	182	16	Q8P3E8	Q8p3e8 xanthomonas
679	6	2.2	169	2	Q7WVP5	Q7wvp5 uncultured	752	6	2.2	182	16	Q8NUE8	Q8nue8 staphylococ
680	6	2.2	169	2	Q7WVP3	Q7wvp3 uncultured	753	6	2.2	183	5	Q26747	Q26747 trypanosoma
681	6	2.2	169	2	Q7WVP2	Q7wvp2 uncultured	754	6	2.2	183	16	Q83C93	Q83c93 coxiella bu
682	6	2.2	169	2	Q7WVP2	Q7wvp2 uncultured	755	6	2.2	184	15	Q8DCX9	Q8dcx9 vibrio vuln
683	6	2.2	169	2	Q7WYN8	Q7wyn8 uncultured	756	6	2.2	184	17	Q8TSV7	Q8tsv7 methanosarc
684	6	2.2	169	2	Q7WYN7	Q7wyn7 uncultured	757	6	2.2	185	2	Q93L86	Q93l86 burkholderi
685	6	2.2	169	2	Q7WYN6	Q7wyn6 uncultured	758	6	2.2	185	16	Q9PD04	Q9pd04 xylella fas
686	6	2.2	169	2	Q7WYN5	Q7wyn5 uncultured	759	6	2.2	185	16	Q8PFM4	Q8pfm4 xanthomonas
687	6	2.2	169	2	Q7WYN4	Q7wyn4 uncultured	760	6	2.2	185	16	Q83E25	Q83e25 coxiella bu
688	6	2.2	169	2	Q7WYN1	Q7wyn1 uncultured	761	6	2.2	186	2	Q9AP03	Q9ap03 burkholderi
689	6	2.2	169	2	Q9BMJ9	Q9bmj9 homo sapien	762	6	2.2	186	15	Q8NMS8	Q8nms8 corynebacte
690	6	2.2	169	2	Q9BMD9	Q9bmd9 homo sapien	763	6	2.2	186	16	Q8DSE6	Q8dse6 vibrio vuln
691	6	2.2	169	10	Q7Y079	Q7y079 dactisca glo	764	6	2.2	187	2	Q54805	Q54805 streptococc
692	6	2.2	170	2	Q9RHS7	Q9rhs7 lactococcus	765	6	2.2	187	16	Q7VKI3	Q7vki3 haemophilus
693	6	2.2	170	8	Q8MLH6	Q8mlh6 chaetosphae	766	6	2.2	187	16	Q8ZS11	Q8zs11 anabaena sp
694	6	2.2	170	9	Q8SCU5	Q8scu5 pseudomonas	767	6	2.2	188	11	Q921P1	Q921p1 mus musculu
695	6	2.2	170	10	Q94IT6	Q94it6 oryza sativ	768	6	2.2	189	15	Q8DNM5	Q8dnm5 streptococc
696	6	2.2	170	10	Q7KCH7	Q7kch7 oryza sativ	769	6	2.2	190	8	Q7YLP0	Q7ylp0 thuga plica
697	6	2.2	170	16	Q9RUD0	Q9rjd0 deinococcus	770	6	2.2	190	17	Q97Y51	Q97y51 sulfolobus
698	6	2.2	170	16	Q97PY6	Q97py6 streptococc	771	6	2.2	190	17	Q97Y51	Q97y51 caenorhadi
699	6	2.2	170	16	Q8CYL2	Q8cyl2 streptococc	772	6	2.2	191	16	Q899R5	Q899r5 clostridium
700	6	2.2	170	16	Q83J95	Q83j95 shigella fl	773	6	2.2	191	16	Q9HW07	Q9hw07 pseudomonas
701	6	2.2	171	2	Q52355	Q52355 mycoplasma	774	6	2.2	192	15	Q8N0F6	Q8n0f6 plasmodium
702	6	2.2	171	16	P73536	P73536 synechocyst	775	6	2.2	193	5	Q8S1P6	Q8s1p6 oryza sativ
703	6	2.2	171	16	Q8CUC3	Q8cuc3 staphylococ	776	6	2.2	193	15	Q8DQK2	Q8dqk2 human immun
704	6	2.2	171	16	Q81RK1	Q81rk1 bacillus an	777	6	2.2	193	16	Q82GQ3	Q82gq3 streptomyce
705	6	2.2	173	16	Q9J5J2	Q9j5j2 chlamydia p	778	6	2.2	193	16	Q7YX72	Q7yx72 mycobacteri
706	6	2.2	173	16	Q87WN6	Q87wn6 pseudomonas	779	6	2.2	194	8	Q559M1	Q559m1 ligia occid
707	6	2.2	173	17	Q8TNN8	Q8tnn8 methanosarc	780	6	2.2	194	16	Q9RUX7	Q9rux7 deinococcus
708	6	2.2	174	2	Q8KSV6	Q8ksv6 methylocald	781	6	2.2	195	4	Q56H61	Q56h61 homo sapien
709	6	2.2	174	2	Q8KS77	Q8ks77 methylocald	782	6	2.2	195	15	Q9RZM6	Q9rzm6 deinococcus
710	6	2.2	174	2	Q8KT04	Q8kt04 methylocald	783	6	2.2	195	16	Q8KCS5	Q8kcs5 chlorobium
711	6	2.2	174	2	Q8KT03	Q8kt03 methylocald	784	6	2.2	196	10	Q7XKX3	Q7xkx3 oryza sativ
712	6	2.2	174	6	Q97625	Q97625 bos taurus	785	6	2.2	196	16	Q8U7Q6	Q8u7q6 agrobacteri
713	6	2.2	175	10	Q9M8J6	Q9m8j6 arabidopsis	786	6	2.2	196	16	Q88LV1	Q88lv1 pseudomonas
714	6	2.2	175	16	Q8PR21	Q8pr21 xanthomonas	787	6	2.2	196	16	Q7W9S2	Q7w9s2 bordetella
715	6	2.2	176	2	Q54413	Q54413 methylocald	788	6	2.2	197	5	Q9SU01	Q9su01 dictyosteli
716	6	2.2	176	2	Q54411	Q54411 methylocald	789	6	2.2	197	15	Q9F3I9	Q9f3i9 streptomyce
717	6	2.2	176	2	Q54412	Q54412 methylocald	790	6	2.2	197	17	Q9V2N9	Q9v2n9 pyrococcus
718	6	2.2	176	4	Q96BA4	Q96ba4 homo sapien	791	6	2.2	197	17	Q37A17	Q37a17 thermoplasma
719	6	2.2	176	16	Q7VLM3	Q7vlm3 haemophilus	792	6	2.2	198	5	Q9Y3J9	Q9y3j9 plasmodium
720	6	2.2	176	17	Q97AN4	Q97an4 thermoplasma	793	6	2.2	198	8	Q85M15	Q85m15 paraclobacta
721	6	2.2	177	2	Q9ZH26	Q9zh26 pseudomonas	794	6	2.2	198	5	Q77005	Q77005 mesocricetu
722	6	2.2	177	2	Q9S6C0	Q9s6c0 pseudomonas	795	6	2.2	198	11	Q8ZLE8	Q8zle8 salmonella
723	6	2.2	177	15	Q8JF06	Q8jfo6 human immu	796	6	2.2	198	16	Q8XKRO	Q8xkro escherichia
724	6	2.2	177	16	Q8ZQJ8	Q8zqj8 streptomyce	797	6	2.2	198	16	Q8Z2S2	Q8z2s2 salmonella
725	6	2.2	178	15	P88235	P88235 human immu	798	6	2.2	198	16	Q8Z2S2	Q8z2s2 salmonella
726	6	2.2	178	16	Q25673	Q25673 helicobacte	799	6	2.2	198	16	Q81LQ9	Q81lq9 bacillus an
727	6	2.2	178	16	Q9ZM28	Q9zm28 helicobacte	800	6	2.2	198	16	Q83J84	Q83j84 shigella fl
728	6	2.2	178	16	Q8BDT8	Q8bdt8 pseudomonas	801	6	2.2	198	17	Q9YBK5	Q9ybk5 aeropyrum p
729	6	2.2	179	4	Q8TC97	Q8tc97 homo sapien	802	6	2.2	199	4	Q9HAMS	Q9hams homo sapien
730	6	2.2	179	4	Q8WAC6	Q8wac6 ligia exoti	803	6	2.2	199	11	Q9EPH2	Q9eph2 rattus norv
731	6	2.2	179	10	Q9LFI8	Q9lif8 arabidopsis	804	6	2.2	199	16	Q9CIU5	Q9ciu5 pasteurella
732	6	2.2	179	11	Q9D9Y3	Q9dy93 mus musculu	805	6	2.2	200	2	Q9ALU8	Q9alu8 enterobacte
733	6	2.2	179	11	Q9DA64	Q9da64 mus musculu	806	6	2.2	200	5	Q817Z1	Q817z1 metapneus
734	6	2.2	179	16	Q8FLE7	Q8fle7 corynebacte	807	6	2.2	200	16	Q92212	Q92212 thizobium m
735	6	2.2	179	16	Q8YF00	Q8yfo0 bradyrhizob	808	6	2.2	200	16	Q8NM52	Q8nm52 corynebacte
736	6	2.2	180	2	Q57025	Q57025 escherichia	809	6	2.2	200	16	Q8FZ71	Q8fz71 bruceella su
737	6	2.2	180	3	Q74882	Q74882 schizosacch	810	6	2.2	200	16	Q82DX3	Q82dx3 streptomyce
738	6	2.2	180	10	Q9LSC7	Q9lsc7 oryza sativ	811	6	2.2	201	3	Q8US91	Q8us91 schizosacch
739	6	2.2	180	16	Q9KSP8	Q9ksp8 bacillus ha	812	6	2.2	201	6	Q8WNV4	Q8wnv4 ovis aries
740	6	2.2	180	16	Q98QJ6	Q98qaj6 mycoplasma	813	6	2.2	201	8	Q9G1D2	Q9g1d2 caridina ze
741	6	2.2	181	2	Q9BMV7	Q9bmrv7 rhodobacter	814	6	2.2	201	8	Q9G1D3	Q9g1d3 caridina ze
742	6	2.2	181	2	Q9ETV6	Q9etv6 corynebacte	815	6	2.2	201	8	Q9G1D3	Q9g1d3 caridina ze
743	6	2.2	181	4	Q14608	Q14608 homo sapien	816	6	2.2	201	8	Q9G469	Q9g469 caridina ze
744	6	2.2	181	4	Q9BZM4	Q9bzm4 homo sapien	817	6	2.2	201	8	Q9G469	Q9g469 caridina ze
745	6	2.2	181	16	Q9KPE7	Q9kpe7 vibrio chol	818	6	2.2	201	17	Q29531	Q29531 archaeoglob
746	6	2.2	181	16	Q9X826	Q9x826 streptomyce	819	6	2.2	202	2	Q8KVT1	Q8kvt1 staphylococ

820	6	2.2	202	2	09KWM2	09KWM2 staphylococ	892	6	2.2	219	5	09VJR3	09VJR3 drosophila
821	6	2.2	202	2	053757	053757 staphylococ	894	6	2.2	219	5	08IDV4	08IDV4 plasmodium
822	6	2.2	202	5	08NOG1	08NOG1 plasmodium	895	6	2.2	219	8	08MB42	08MB42 bnto panthe
823	6	2.2	202	16	097LE7	097LE7 clostridium	896	6	2.2	219	8	08W7U5	08W7U5 ligia perki
824	6	2.2	202	16	08DLE8	08DLE8 synchococ	897	6	2.2	219	8	08WAB8	08WAB8 ligia hawai
825	6	2.2	202	16	07VGM4	07VGM4 helicobacte	898	6	2.2	219	8	08WAB5	08WAB5 ligia hawai
826	6	2.2	203	5	09UYI2	09UYI2 caenorhabd	899	6	2.2	219	8	08WAC2	08WAC2 ligia hawai
827	6	2.2	203	5	045289	045289 caenorhabd	900	6	2.2	219	8	08WAC1	08WAC1 ligia hawai
828	6	2.2	203	13	090XY0	090XY0 fungu rubrip	901	6	2.2	219	10	09SP76	09SP76 gossypium a
829	6	2.2	203	16	08KXND	08KXND clostridium	902	6	2.2	219	10	09S7S9	09S7S9 gossypium 1
830	6	2.2	203	16	08CUD5	08CUD5 staphylococ	903	6	2.2	219	16	09PBR3	09PBR3 xylella fas
831	6	2.2	203	16	08S9RS	08S9RS bradyrhizob	904	6	2.2	219	16	07W329	07W329 bordetella
832	6	2.2	203	16	083L89	083L89 shigella fl	905	6	2.2	220	16	09S2X4	09S2X4 streptomyce
833	6	2.2	204	16	08YIP0	08YIP0 bruceella me	906	6	2.2	220	16	089ZS6	089ZS6 bacteroides
834	6	2.2	205	2	08GKM7	08GKM7 streptococ	907	6	2.2	221	9	021979	021979 bacterioph
835	6	2.2	205	16	051738	051738 borrelia bu	908	6	2.2	221	10	08VY90	08VY90 arabidopsi
836	6	2.2	205	16	08BRK3	08BRK3 pseudomonas	909	6	2.2	221	10	08LD04	08LD04 arabidopsi
837	6	2.2	205	16	07U7T8	07U7T8 synchococ	910	6	2.2	221	16	08XMR6	08XMR6 ralsconia s
838	6	2.2	206	2	044801	044801 borrelia bu	911	6	2.2	221	17	097UD7	097UD7 sulfolobus
839	6	2.2	206	8	08MAB7	08MAB7 ligia hawai	912	6	2.2	222	2	P72322	P72322 rhodospirill
840	6	2.2	206	8	08WACO	08WACO ligia hawai	913	6	2.2	223	16	08FEW9	08FEW9 fusobacteri
841	6	2.2	206	8	08WAB9	08WAB9 ligia hawai	914	6	2.2	223	16	08FM15	08FM15 corynebacte
842	6	2.2	206	8	08W7H2	08W7H2 ligia hawai	915	6	2.2	223	16	08E3S8	08E3S8 streptococ
843	6	2.2	206	16	08YER8	08YER8 bruceella me	916	6	2.2	223	16	08DYC2	08DYC2 streptococ
844	6	2.2	206	16	08G317	08G317 bruceella su	917	6	2.2	223	16	08B1Y6	08B1Y6 lactobacill
845	6	2.2	206	16	087S17	087S17 vibrio para	918	6	2.2	223	16	08ZSK5	08ZSK5 streptomyce
846	6	2.2	207	5	08N0G0	08N0G0 plasmodium	919	6	2.2	223	16	07UEA5	07UEA5 rhodopirell
847	6	2.2	207	5	016575	016575 caenorhabd	920	6	2.2	224	2	09AF13	09AF13 shigella fl
848	6	2.2	207	10	09LEF1	09LEF1 arabidopsi	921	6	2.2	224	2	08V1H8	08V1H8 shigella fl
849	6	2.2	207	16	07UUY0	07UUY0 rhodopirell	922	6	2.2	224	10	09FHS3	09FHS3 arabidopsi
850	6	2.2	208	2	0849C2	0849C2 haemophilus	923	6	2.2	224	16	09BR12	09BR12 rhizobium 1
851	6	2.2	208	8	085MR8	085MR8 ephemerella	924	6	2.2	224	16	08G2U7	08G2U7 bruceella su
852	6	2.2	208	8	085M07	085M07 periplaneta	925	6	2.2	224	16	081ZQ7	081ZQ7 shigella fl
853	6	2.2	208	8	085M19	085M19 chrysotoxum	926	6	2.2	224	16	09AFN6	09AFN6 shigella fl
854	6	2.2	208	8	085M07	085M07 eurylophell	927	6	2.2	224	16	07UD61	07UD61 shigella fl
855	6	2.2	208	16	08NMQ0	08NMQ0 corynebacte	928	6	2.2	225	16	089US5	089US5 bradyrhizob
856	6	2.2	208	16	08BET3	08BET3 shewanella	929	6	2.2	226	16	08Y271	08Y271 anabaena sp
857	6	2.2	208	16	07MCY6	07MCY6 bordetella	930	6	2.2	226	16	08BMT1	08BMT1 pseudomonas
858	6	2.2	209	2	09FBW4	09FBW4 bordetella	931	6	2.2	226	16	08BMT1	08BMT1 streptomyce
859	6	2.2	209	5	021621	021621 caenorhabd	932	6	2.2	227	2	082YF5	082YF5 synchococ
860	6	2.2	209	16	09HX40	09HX40 pseudomonas	933	6	2.2	227	2	054768	054768 streptococ
861	6	2.2	209	16	08NNU2	08NNU2 corynebacte	934	6	2.2	227	5	08K1P0	08K1P0 rhizobium 1
862	6	2.2	209	16	08A5E5	08A5E5 bacteroides	935	6	2.2	227	5	086DP6	086DP6 sepiia offic
863	6	2.2	210	5	0966C8	0966C8 caenorhabd	936	6	2.2	227	5	0869N5	0869N5 dictyosteli
864	6	2.2	210	16	09RTF2	09RTF2 deinococcus	937	6	2.2	227	16	08ZJU0	08ZJU0 salmonella
865	6	2.2	210	16	056649	056649 synchocyst	938	6	2.2	227	16	08ZOU8	08ZOU8 salmonella
866	6	2.2	210	16	09RG61	09RG61 pseudomonas	939	6	2.2	227	17	08TU53	08TU53 methanopyru
867	6	2.2	210	16	08F1E7	08F1E7 leptospira	940	6	2.2	228	5	081H06	081H06 drosophila
868	6	2.2	210	16	07UID6	07UID6 rhodopirell	941	6	2.2	228	10	08L9U4	08L9U4 arabidopsi
869	6	2.2	211	5	081RH2	081RH2 drosophila	942	6	2.2	228	10	09FCM1	09FCM1 zea mays (m
870	6	2.2	211	16	08PBC2	08PBC2 xanthomonas	943	6	2.2	228	10	09M1R7	09M1R7 arabidopsi
871	6	2.2	211	16	0896B8	0896B8 clostridium	944	6	2.2	228	16	0837E0	0837E0 mycobacteri
872	6	2.2	213	10	08VY51	08VY51 arabidopsi	945	6	2.2	228	16	087EWS	087EWS vibrio para
873	6	2.2	213	16	099YH3	099YH3 streptococ	946	6	2.2	228	17	07UIQ2	07UIQ2 mycobacteri
874	6	2.2	213	16	099W74	099W74 staphylococ	947	6	2.2	228	16	P94946	P94946 methanopyru
875	6	2.2	213	16	08P0C8	08P0C8 streptococ	948	6	2.2	230	13	07SVY9	07SVY9 brachydanio
876	6	2.2	213	16	08PPR9	08PPR9 streptococ	949	6	2.2	230	16	08NPI4	08NPI4 corynebacte
877	6	2.2	214	16	092W95	092W95 rhizobium m	950	6	2.2	230	16	08BQC0	08BQC0 pseudomonas
878	6	2.2	214	16	08E7J5	08E7J5 streptococ	951	6	2.2	231	2	045961	045961 clostridium
879	6	2.2	214	16	08E1L9	08E1L9 streptococ	952	6	2.2	231	2	059113	059113 burkholderi
880	6	2.2	215	6	0810J1	0810J1 bos taurus	953	6	2.2	231	11	08CF79	08CF79 mus musculu
881	6	2.2	215	10	09M3U5	09M3U5 nicotiana t	954	6	2.2	231	16	088QB7	088QB7 pseudomonas
882	6	2.2	215	16	0879E8	0879E8 streptococ	955	6	2.2	231	17	09Y9D7	09Y9D7 aeropyrum p
883	6	2.2	215	17	08PW73	08PW73 methanosarc	956	6	2.2	231	17	0978V6	0978V6 thermoplasm
884	6	2.2	216	2	08K1Y4	08K1Y4 uncultured	957	6	2.2	232	2	07MWV4	07MWV4 alcaligenes
885	6	2.2	216	2	09F243	09F243 actinobacill	958	6	2.2	232	5	09VUC1	09VUC1 drosophila
886	6	2.2	216	16	07WVJ5	07WVJ5 bordetella	959	6	2.2	232	16	08XW02	08XW02 ralsconia s
887	6	2.2	216	16	07VZ47	07VZ47 bordetella	960	6	2.2	232	16	08DTB6	08DTB6 streptococ
888	6	2.2	217	8	08W6U6	08W6U6 pteris hiru	961	6	2.2	232	16	07UC12	07UC12 shigella fl
889	6	2.2	217	16	08DQL8	08DQL8 streptococ	962	6	2.2	233	2	09MJ90	09MJ90 actinobacill
890	6	2.2	218	8	0961J2	0961J2 homo sapien	963	6	2.2	233	16	08XZN3	08XZN3 ralsconia s
891	6	2.2	218	8	08WAB6	08WAB6 ligia perki	964	6	2.2	233	16	08DSX2	08DSX2 vibrio vuln
892	6	2.2	218	17	08THJ3	08THJ3 methanosarc	965	6	2.2	234	16	09BLF6	09BLF6 rhizobium 1

```

966 6 2.2 235 16 09CJ22 pasteurella
967 6 2.2 236 2 08KMA1 08Kw1 ruegeria sp
968 6 2.2 236 12 08Q575 08Q575 chimpanzee
969 6 2.2 236 16 08NMM0 08NMM0 corynebacter
970 6 2.2 237 16 08AL65 08AL65 pseudomons
971 6 2.2 237 2 09AMC6 09AMC6 streptococ
972 6 2.2 237 2 08KSP9 08KSP9 streptomyc
973 6 2.2 237 2 0848N0 0848N0 gamma-prote
974 6 2.2 237 2 0848M0 0848M0 gamma-prote
975 6 2.2 237 16 09KXM4 09KXM4 vibrio chol
976 6 2.2 237 16 097RU0 097RU0 streptococc
977 6 2.2 237 16 08CZ07 08CZ07 streptococc
978 6 2.2 238 2 048988 048988 mycoplasma
979 6 2.2 238 2 08KX00 08KX00 desulfobacu
980 6 2.2 238 16 089PX3 089PX3 bradyrhizob
981 6 2.2 239 10 09LFJ3 09LFJ3 arabidopsis
982 6 2.2 239 16 092NZ4 092NZ4 rhizobium m
983 6 2.2 239 16 08UDK5 08UDK5 agrobacteri
984 6 2.2 239 17 097WM0 097WM0 sulfolobus
985 6 2.2 239 17 08TPX2 08TPX2 methanosa
986 6 2.2 239 17 08PTU3 08PTU3 methanosa
987 6 2.2 240 2 093LU0 093LU0 streptococ
988 6 2.2 240 11 099UX2 099UX2 mus musculi
989 6 2.2 240 16 08XBD7 08XBD7 escherichia
990 6 2.2 240 16 08FF80 08FF80 escherichia
991 6 2.2 240 16 083K55 083K55 shigella fl
992 6 2.2 241 5 09UYO0 09UYO0 leishmania
993 6 2.2 241 16 08ZSN1 08ZSN1 salmonella
994 6 2.2 241 16 08ZSU7 08ZSU7 nitrosomona
995 6 2.2 242 3 09P8A6 09P8A6 agarticus bi
996 6 2.2 242 10 09F155 09F155 arabidopsis
997 6 2.2 242 16 092PFD0 092PFD0 listeria in
998 6 2.2 242 16 08YAJ6 08YAJ6 listeria mo
999 6 2.2 242 16 08SIT9 08SIT9 streptomyc
1000 6 2.2 242 16 08BPF0 08BPF0 pseudomons

```

ALIGNMENTS

RESULT 1

```

ID 07WM93 PRELIMINARY; PRT; 270 AA.
AC 07WM93;
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE 28kDa outer membrane lipoprotein.
GN HLRP OR HD1100.
OS Haemophilus ducreyi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_Taxid=730;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3500HP / ATCC 700724;
RA Munson R.S. Jr., Ray W.C., Mahairas G., Sabo P., Murgur R.,
RA Johnson L., Nguyen D., Wang J., Forst C., Hood L.;
RT "The complete genome sequence of Haemophilus ducreyi."
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017153; AAP95966.1; -
KW Lipoprotein, Complete proteome.
SQ SEQUENCE 270 AA; 29535 MW; FC077F10B13304AD CRC64;

```

Query Match 5.8%; Score 16; DB 16; Length 270;
 Best Local Similarity 100.0%; Pred. No. 1.6e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 KDSEKGLNLYVGN 115
 DB 95 KDSEKGLNLYVGN 110

RESULT 2

```

ID 083MC6 PRELIMINARY; PRT; 271 AA.
AC 083MC6;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Putative lipoprotein.
GN YABC OR SF0188 OR S0190.
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_Taxid=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=301 / Serotype 2a;
RA MEDLINE=2272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157."
RL Nucleic Acids Res. 30:4432-4441(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RA MEDLINE=22590774; PubMed=12704152;
RA Wei J., Goldberg M.B., Burdard V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blatner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T."
RL Infect. Immun. 71:2775-2786(2003).
RN [3]
RP EMBL; AE015055; AAN41850.1; -
DR EMBL; AE016978; AAP15730.1; -
DR InterPro: IPR004872; Lipoprotein_9.
DR InterPro: IPR000437; Prok_Lipoprot_S.
DR InterPro: IPR004478; YaeC.
DR Pfam: PF03180; Lipoprotein_9; 1.
DR TIGRFAMs; TIGR00363; TIGR00363; 1.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Lipoprotein; Complete proteome.
SQ SEQUENCE 271 AA; 29417 MW; C3D7D01F85CD2BFE CRC64;

```

Query Match 4.0%; Score 11; DB 16; Length 271;
 Best Local Similarity 100.0%; Pred. No. 0.021;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 224 FVEDKDSPPVN 234
 DB 219 FVEDKDSPPVN 229

RESULT 3

```

ID 07WIF5 PRELIMINARY; PRT; 271 AA.
AC 07WIF5;
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Putative exported protein.
GN BB2896.
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_Taxid=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RB50 / ATCC BAA-588;
RA MEDLINE=22827954; PubMed=12910271;

```

RA Parkhill J., Sebathia M., Preston A., Murphy L.D., Thomson N.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Feltwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagsels K.,
 RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
 RA Unwin L., Whitehead S., Barrall B.G., Maskell D.J.,
 RT "Comparative analysis of the genome sequences of *Bordetella pertussis*,
 RT *Bordetella parapertussis* and *Bordetella bronchiseptica*.";
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL; BX640445; CAE33388.1; -.
 KM Complete proteome.
 SQ SEQUENCE 271 AA; 29643 MW; FC5B8C19A837AB0E CRC64;

Query Match 3.6%; Score 10; DB 16; Length 271;
 Best Local Similarity 100.0%; Pred. No. 0.22;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 GATTAVPNDP 146
 DB 127 GATTAVPNDP 136

RESULT 4
 QYVYM6 PRELIMINARY; PRT; 271 AA.

DT 01-OCT-2003 (TREMBLrel. 25, Created)
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Putative exported protein.
 GN BP1292.

OS *Bordetella pertussis*.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; *Bordetella*.
 NC NCB1_TaxID=520;
 RX [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
 RX MEDLINE=22827954; PubMed=12910271;

RA Parkhill J., Sebathia M., Preston A., Murphy L.D., Thomson N.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Feltwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagsels K.,
 RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
 RA Unwin L., Whitehead S., Barrall B.G., Maskell D.J.,
 RT "Comparative analysis of the genome sequences of *Bordetella pertussis*,
 RT *Bordetella parapertussis* and *Bordetella bronchiseptica*.";
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL; BX640445; CAE33388.1; -.
 KM Complete proteome.
 SQ SEQUENCE 271 AA; 29614 MW; E75B8C1C2337AB06 CRC64;

Query Match 3.6%; Score 10; DB 16; Length 271;
 Best Local Similarity 100.0%; Pred. No. 0.22;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 GATTAVPNDP 146
 DB 127 GATTAVPNDP 136

RESULT 5
 ID Q7M6I3 PRELIMINARY; PRT; 281 AA.
 AC Q7M6I3;

DT 01-OCT-2003 (TREMBLrel. 25, Created)
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Putative exported protein.
 GN BP2926.
 OS *Bordetella parapertussis*.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; *Bordetella*.
 NC NCB1_TaxID=519;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=12822 / ATCC BAA-587;
 RX MEDLINE=22827954; PubMed=12910271;
 RA Parkhill J., Sebathia M., Preston A., Murphy L.D., Thomson N.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Quail M.A.,
 RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Feltwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagsels K.,
 RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
 RA Unwin L., Whitehead S., Barrall B.G., Maskell D.J.,
 RT "Comparative analysis of the genome sequences of *Bordetella pertussis*,
 RT *Bordetella parapertussis* and *Bordetella bronchiseptica*.";
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL; BX640432; CAE38219.1; -.
 KM Complete proteome.
 SQ SEQUENCE 281 AA; 30591 MW; C97FA451E40E3315 CRC64;

Query Match 3.6%; Score 10; DB 16; Length 281;
 Best Local Similarity 100.0%; Pred. No. 0.23;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 GATTAVPNDP 146
 DB 137 GATTAVPNDP 146

RESULT 6
 ID Q69441 PRELIMINARY; PRT; 259 AA.
 AC Q69441;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE 29 kDa immunogenic protein.
 OS *Legionella pneumophila*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
 OC Legionellaceae; *Legionella*.
 NC NCB1_TaxID=446;
 RX [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=Corby;
 RA Stendel C., Helbig J., Lueck C.,
 RT "Cloning and antigenic characterization of a 29 kDa protein from
 RT *Legionella pneumophila* strain Corby.";
 RL Submitted (APR-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL; A005668; CA06664.1; -.
 DR InterPro; IPR004872; Lipoprotein_9.
 DR InterPro; IPR000457; Prok_LipoProt_S.
 DR InterPro; IPR004478; YaeC.
 DR Pfam; PFO3180; Lipoprotein_9; 1.
 DR TIGRPFAMS; TIGR00363; TIGR00363; 1.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 SQ SEQUENCE 259 AA; 28614 MW; 17FDB7AF6C4ADA66 CRC64;

Query Match 3.3%; Score 9; DB 2; Length 259;
 Best Local Similarity 100.0%; Pred. No. 2.3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 IAVNDPSN 148
 ID IAVNDPSN 148

Db 125 IAVNPDSN 133

RESULT 7

Q7WEM2 PRELIMINARY; PRT; 262 AA.

AC Q7WEM2 01-OCT-2003 (TREMBlrel. 25, Created)

DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)

DE Putative exported protein.

GN BB4519.

OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).

OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;

OC Alcaligenaceae; Bordetella.

OX NCBI_TaxID=518;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=RB50 / ATCC BAA-588;

RX MEDLINE=22827954; PubMed=12910271.

RA Parkhill J., Sebald M., Preston A., Murphy L.D., Thomson N.,

RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,

RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,

RA Achman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,

RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,

RA Felwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,

RA Leather S., Moul S., Norberczak H., O'Neill S., Ormond D., Price C.,

RA Rabinowitsch E., Rutter S., Sanders M., Saunders R., Seeger K.,

RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,

RA Unwin L., Whitehead S., Barrrell B.G., Maskell D.J.,

RT "Comparative analysis of the genome sequences of Bordetella pertussis,

RT Bordetella parapertussis and Bordetella bronchiseptica."

RL Nat. Genet. 35:32-40(2003).

DR EMBL: BX640450; CAB34882.1; -

KM Complete proteome.

SQ SEQUENCE 262 AA; 28490 MW; 0DC80B1739FF4663 CRC64;

Query Match 3.3%; Score 9; DB 16; Length 262;

Best Local Similarity 100.0%; Pred. No. 2.3;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 133 ELKDGATTA 141

Db 119 ELKDGATTA 127

RESULT 8

Q7W317 PRELIMINARY; PRT; 262 AA.

AC Q7W317 01-OCT-2003 (TREMBlrel. 25, Created)

DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)

DE Putative exported protein.

GN BP4046.

OS Bordetella parapertussis.

OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;

OC Alcaligenaceae; Bordetella.

OX NCBI_TaxID=519;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=12822 / ATCC BAA-587;

RX MEDLINE=22827954; PubMed=12910271;

RA Parkhill J., Sebald M., Preston A., Murphy L.D., Thomson N.,

RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,

RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,

RA Achman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,

RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,

RA Felwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,

RA Leather S., Moul S., Norberczak H., O'Neill S., Ormond D., Price C.,

RA Rabinowitsch E., Rutter S., Sanders M., Saunders R., Seeger K.,

RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,

RA Unwin L., Whitehead S., Barrrell B.G., Maskell D.J.,

RT "Comparative analysis of the genome sequences of Bordetella pertussis,

RT Bordetella parapertussis and Bordetella bronchiseptica."

RL Nat. Genet. 35:32-40(2003).

DR EMBL: BX640412; CAB44716.1; -

KM Complete proteome.

SQ SEQUENCE 262 AA; 28524 MW; E0A56B1A59FF4EC9 CRC64;

Query Match 3.3%; Score 9; DB 16; Length 262;

Best Local Similarity 100.0%; Pred. No. 2.3;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 133 ELKDGATTA 141

Db 119 ELKDGATTA 127

RT "Comparative analysis of the genome sequences of Bordetella pertussis,

RT Bordetella parapertussis and Bordetella bronchiseptica."

RL Nat. Genet. 35:32-40(2003).

DR EMBL: BX640435; CAB39329.1; -

KM Complete proteome.

SQ SEQUENCE 262 AA; 28506 MW; D4C18652307EBB2A CRC64;

Query Match 3.3%; Score 9; DB 16; Length 262;

Best Local Similarity 100.0%; Pred. No. 2.3;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 133 ELKDGATTA 141

Db 119 ELKDGATTA 127

RESULT 9

Q7VSM1 PRELIMINARY; PRT; 262 AA.

AC Q7VSM1 01-OCT-2003 (TREMBlrel. 25, Created)

DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)

DE Putative exported protein.

GN BP0385.

OS Bordetella pertussis.

OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;

OC Alcaligenaceae; Bordetella.

OX NCBI_TaxID=520;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;

RX MEDLINE=22827954; PubMed=12910271.

RA Parkhill J., Sebald M., Preston A., Murphy L.D., Thomson N.,

RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,

RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,

RA Achman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,

RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,

RA Felwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,

RA Leather S., Moul S., Norberczak H., O'Neill S., Ormond D., Price C.,

RA Rabinowitsch E., Rutter S., Sanders M., Saunders R., Seeger K.,

RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,

RA Unwin L., Whitehead S., Barrrell B.G., Maskell D.J.,

RT "Comparative analysis of the genome sequences of Bordetella pertussis,

RT Bordetella parapertussis and Bordetella bronchiseptica."

RL Nat. Genet. 35:32-40(2003).

DR EMBL: BX640412; CAB44716.1; -

KM Complete proteome.

SQ SEQUENCE 262 AA; 28524 MW; E0A56B1A59FF4EC9 CRC64;

Query Match 3.3%; Score 9; DB 16; Length 262;

Best Local Similarity 100.0%; Pred. No. 2.3;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 133 ELKDGATTA 141

Db 119 ELKDGATTA 127

RESULT 10

Q8YA74 PRELIMINARY; PRT; 273 AA.

AC Q8YA74 01-MAR-2002 (TREMBlrel. 20, Created)

DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)

DE Putative lipoprotein.

GN LMO285.

OS Listeria monocytogenes.

OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.

OC NCBI_TaxID=1639;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ESD-e / Serovar 1/2a;
 RX MEDLINE=21537279; PubMed=11679669;
 RA Clauser P., Frangeul L., Buchrieser C., Ruenick C., Amend A.,
 RA Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
 RA Charbit A., Chebouni F., Couve E., de Daruvar A., Deloux P.,
 RA Dammann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
 RA Eutican K.-D., Feihl H., Garcia-del Portillo F., Garrido P.,
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
 RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkac G.,
 RA Madueno E., Maitouran A., Mata Vicente U., Ng E., Negari H.,
 RA Nordstrek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
 RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
 RA Vazquez-Boland U.-A., Voss H., Weiland J., Cossart P.,
 RT "Comparative genomics of *Listeria species*,"
 RL Science 294:849-852(2001).
 RL EMBL; AF591974; CAD00812.1; -
 DR PIR; AF1110; AF1110.
 DR Listlist; LMO00285; -
 DR InterPro; IPR004872; Lipoprotein_9.
 DR InterPro; IPR000437; Prok_lipoprot_S.
 DR InterPro; IPR004478; Yaec-
 DR Pfam; PF03180; Lipoprotein_9; 1.
 DR TIGRPFAM; TIGR00363; TIGR00363; 1.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 KW Lipoprotein; Hypothetical protein; Complete proteome.
 SQ SEQUENCE 273 AA; 29811 MW; AC874D27BBA3AABC CRC64;

Query Match 3.3%; Score 9; DB 16; Length 273;
 Best Local Similarity 100.0%; Pred. No. 2.4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 IAVNPDPN 148
 Db 136 IAVNPDPN 144

RESULT 11
 Q7WML8 PRELIMINARY; PRT; 276 AA.
 AC Q7WML8;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Putative transglycosylase.
 GN Bbl373.
 OS Bordetella bronchiseptica (Alcaligenes bronchiseptus).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 OX NCBI_TaxID=518;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RS50 / ATCC BAA-588;
 RX MEDLINE=22827954; PubMed=12910271;
 RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Feltham T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagsels K.,
 RA Rabinowitsch E., Rutter S., Skelton J., Squares R., Squares S., Stevens K.,
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
 RA Umwin L., Whitehead S., Barrett B.G., Maskell D.J.,
 RT "Comparative analysis of the genome sequences of *Bordetella pertussis*,
 RT *Bordetella parapertussis* and *Bordetella bronchiseptica*,"
 RL Nat. Genet. 35:32-40(2003).
 RL EMBL; BX640441; CAB31871.1; -
 KM Complete proteome.
 SQ SEQUENCE 276 AA; 28845 MW; C42904282B560533 CRC64;

Query Match 3.3%; Score 9; DB 16; Length 276;
 Best Local Similarity 100.0%; Pred. No. 2.4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ALASGIALA 19
 Db 266 ALASGIALA 274

RESULT 12
 Q7WBS1 PRELIMINARY; PRT; 276 AA.
 AC Q7WBS1;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Putative transglycosylase.
 GN Bpl157.
 OS Bordetella parapertussis.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 OX NCBI_TaxID=519;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=12822 / ATCC BAA-587;
 RX MEDLINE=22827954; PubMed=12910271;
 RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Feltham T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagsels K.,
 RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
 RA Umwin L., Whitehead S., Barrett B.G., Maskell D.J.,
 RT "Comparative analysis of the genome sequences of *Bordetella pertussis*,
 RT *Bordetella parapertussis* and *Bordetella bronchiseptica*,"
 RL Nat. Genet. 35:32-40(2003).
 RL EMBL; BX640426; CAB36458.1; -
 KM Complete proteome.
 SQ SEQUENCE 276 AA; 28845 MW; C42904282B560533 CRC64;

Query Match 3.3%; Score 9; DB 16; Length 276;
 Best Local Similarity 100.0%; Pred. No. 2.4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ALASGIALA 19
 Db 266 ALASGIALA 274

RESULT 13
 Q7VZ71 PRELIMINARY; PRT; 276 AA.
 AC Q7VZ71;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Putative transglycosylase.
 GN Bpl061.
 OS Bordetella pertussis.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 OX NCBI_TaxID=520;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Robama I / ATCC BAA-589 / NCTC 13251;
 RX MEDLINE=22827954; PubMed=12910271;
 RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Feltham T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagsels K.,
 RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,

RA Rabbihowitz E., Rutter S., Sanders M., Saunders D., Seeger K.,
 RA Sharp S., Simmonds M., Skelton U., Squares R., Squares S., Stevens K.,
 RA Unwin L., Whitehead S., Barrett B.G., Maskell D.J.,
 RT "Comparative analysis of the genome sequences of *Bordetella pertussis*,
 RT *Bordetella parapertussis* and *Bordetella bronchiseptica*,"
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL; BX640414; CAB31360.1; -.
 KW Complete proteome.
 SQ SEQUENCE 276 AA; 28879 MW; C0126189C59B362 CRC64;

Query Match 3.3%; Score 9; DB 16; Length 276;
 Best Local Similarity 100.0%; Pred. No. 2.4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 ALASGIALA 19
 |||||
 Db 266 ALASGIALA 274

RESULT 14
 08ID39 PRELIMINARY; PRT; 278 AA.
 AC 08YD39;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE ABC transporter substrate binding protein.
 GN EMBL0338.
 OS *Brucella melitensis*.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Brucellaceae; Brucella.
 CX NCBI_TaxID=29459;
 RN [1]

SEQUENCE FROM N.A.
 RC STRAIN=16M / ATCC 23456 / Biotype 1;
 RX MEDLINE=20020109; PubMed=11756688;
 RA Delvecchio V.G., Kapral V., Redkar R.J., Patra G., Mijer C., Los T.,
 RA Ivanova N., Anderson I., Bhattacharya A., Lykdis A., Reznik G.,
 RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,
 RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
 RA Haselcorn R., Kyriades N., Overbeek R.,
 RT "The genome sequence of the facultative intracellular pathogen
 RT *Brucella melitensis*,"
 RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
 DR EMBL; AEO09671; AALS5580.1; -.
 DR PIR; A13551; A13551.
 DR InterPro; IPR004872; Lipoprotein_9.
 DR InterPro; IPR006311; Tat.
 DR InterPro; IPR004478; YaeC.
 DR Pfam; PF03180; Lipoprotein_9; 1.
 DR TIGRFAMs; TIGR01409; TAT signal seq; 1.
 DR TIGRFAMs; TIGR00363; TIGR00363; 1.
 KW Complete proteome.
 SQ SEQUENCE 278 AA; 30606 MW; C0126189C59B362 CRC64;

Query Match 3.3%; Score 9; DB 16; Length 278;
 Best Local Similarity 100.0%; Pred. No. 2.4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 176 DIVENPKKL 184
 |||||
 Db 177 DIVENPKKL 185

RESULT 15
 08FV86 PRELIMINARY; PRT; 278 AA.
 AC 08FV86;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Lipoprotein, YaeC family.
 GN BRA0960.

OS *Brucella suis*.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Brucellaceae; Brucella.
 CX NCBI_TaxID=29461;
 RN [1]

SEQUENCE FROM N.A.
 RC STRAIN=1330 / Biovar 1;
 RX MEDLINE=22247741; PubMed=12271122;
 RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
 RA Read T.D., Dodson R.J., Umeyam L., Brinkac L.M., Beaman M.J.,
 RA Daugherty S.C., Debey R.T., Durkin A.S., Kolonay J.F., Madupu R.,
 RA Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J., Van Aken S.E.,
 RA Riedmuller S., Mettelin H., Gill S.R., White O., Salzberg S.L.,
 RA Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.,
 RT "The *Brucella suis* genome reveals fundamental similarities between
 RT animal and plant pathogens and symbionts,"
 RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
 DR EMBL; AB014589; AAN34131.1; -.
 DR PIR; A13551; A13551.
 DR TIGR; BRA0960; -.
 DR InterPro; IPR004872; Lipoprotein_9.
 DR InterPro; IPR006311; Tat.
 DR InterPro; IPR004478; YaeC.
 DR Pfam; PF03180; Lipoprotein_9; 1.
 DR TIGRFAMs; TIGR01409; TAT signal seq; 1.
 DR TIGRFAMs; TIGR00363; TIGR00363; 1.
 KW Complete proteome.
 SQ SEQUENCE 278 AA; 30606 MW; C0126189C59B362 CRC64;

Query Match 3.3%; Score 9; DB 16; Length 278;
 Best Local Similarity 100.0%; Pred. No. 2.4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 176 DIVENPKKL 184
 |||||
 Db 177 DIVENPKKL 185

RESULT 16
 09UY77 PRELIMINARY; PRT; 89 AA.
 AC 09UY77;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Molycoprotein converting factor, subunit 1 (MOAD).
 GN PYRAB14200 OR PAB3357.
 OS *Pyrococcus abyssi*.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcaceae;
 OC Pyrococcus.
 CX NCBI_TaxID=29292;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GES / Orsay;
 RA Helig R.;
 RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
 RT structure and evolution,"
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ248287; CAB50325.1; -.
 DR PIR; H75053; H75053.
 DR HSP; P30748; IFMO.
 DR GO; GO:0006790; P:sulfur metabolism; IEX.
 DR InterPro; IPR003749; This.
 DR Pfam; PF02597; This; 1.
 KW Complete proteome.
 SQ SEQUENCE 89 AA; 10252 MW; C93780E376876E2C CRC64;

Query Match 2.9%; Score 8; DB 17; Length 89;
 Best Local Similarity 100.0%; Pred. No. 9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 133 ELKDGAT 140
 |||||

Db 23 ELKOGATI 30

RESULT 17

Q9AGI3 PRELIMINARY; PRT; 255 AA.

AC Q9AGI3; (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 NC NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=2137446; PubMed=11461492;
 RA Atora S.K., Banger M., Lory S., Rampal R.;
 RT "A genomic island in Pseudomonas aeruginosa carries the determinants
 of flagellin glycosylation.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9342-9347(2001).
 DR EMBL; AF332547; AK15335.1; -
 DR GO; GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
 DR InterPro; IPR003329; Cytidylyl_trans.
 DR Pfam; PF02348; CTP_transf_3; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 255 AA; 28299 MW; 969DD8BA564D53B4 CRC64;

Query Match 2.9%; Score 8; DB 2; Length 255;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 VAEVAGV 58
 DB 80 VAEVAGV 87

RESULT 18

Q7XSM6 PRELIMINARY; PRT; 259 AA.

AC Q7XSM6; (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 NC NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CF5;
 RA Atora S.K., Wolfgang M.C., Lory S., Rampal R.;
 RT "Sequence polymorphism in the Glycosylation Island and flagellins of
 Pseudomonas aeruginosa.";
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY280453; AAP35723.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 259 AA; 28789 MW; 2B60BD2857EA9038 CRC64;

Query Match 2.9%; Score 8; DB 2; Length 259;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 VAEVAGV 58
 DB 84 VAEVAGV 91

RESULT 19
 Q8XNLT7 PRELIMINARY; PRT; 259 AA.
 ID Q8XNLT7

AC Q8XNLT7;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Probable transcriptional regulator.
 GN CPE0316
 OS Clostridium perfringens.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 NC NCBI_TaxID=1502;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=13 / Type A;
 RX MEDLINE=2164373; PubMed=11792842;
 RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
 RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
 RT "Complete genome sequence of Clostridium perfringens, an anaerobic
 flesh-eater.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).

DR EMBL; AP003165; BAB0022.1; -
 DR GO; GO:0005622; C:intracellular; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR01034; HTH_Deor.
 DR Pfam; PF00455; deor; 1.
 DR SMART; SM00420; HTH_Deor; 1.
 DR PROSITE; PS00894; HTH_Deor_FAMILY; 1.
 KW Complete proteome.
 SQ SEQUENCE 259 AA; 29210 MW; 83A04B3A2E6F1A2E CRC64;

Query Match 2.9%; Score 8; DB 16; Length 259;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 LEKQGLIK 163
 DB 43 LEKQGLIK 50

RESULT 20

Q8U7G0 PRELIMINARY; PRT; 259 AA.

AC Q8U7G0;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Outer membrane lipoprotein.
 GN ATU4489 OR AGR_L_761.
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
 NC NCBI_TaxID=176299;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=21608550; PubMed=11743193;
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitejima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boyce D., Sr.,
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
 RA Kutyavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
 RA Raymond C., Rouse G., Saenphimmachak C., Xu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chunley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nester E.W.;
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
 C58.";
 RL Science 294:2317-2323(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
 RA Gurello B., Goldman B.S., Cao Y., Askenazi M., Hailing C., Mullin L.,

RA Houniel K., Gordon J., Vaudin M., Tatchouk O., Epp A., Liu F.,
 RA Woliam C., Allinger M., Doughy J., Scott C., Lappas C., Markelz B.,
 RA Planagan C., Crowell C., Gerson J., Lomo C., Sear C., Strud G.,
 RA Cleio C., Slater S.,
 RT "Genome sequence of the plant pathogen and biotechnology agent
 RT Agrobacterium tumefaciens C58.",
 RL Science 294:2323-2328(2001).
 DR EMBL: AF0083177; AAI45283.1; -
 DR EMBL: AF0082377; AAK88953.1; -
 DR PIR: AE3108; AE3108.
 DR PIR: G98178; G98178.
 DR InterPro: IPR004872; Lipoprotein_9.
 DR InterPro: IPR004478; YaeC.
 DR Pfam: PF03180; Lipoprotein_9; 1.
 DR TrEMBL: TIGR00363; TIGR00363; 1.
 KM TrEMBL: TIGR00363; TIGR00363; 1.
 SQ SEQUENCE 259 AA; 27934 MW; 77AD12ED13960861A CRC64;

Query Match 2.9%; Score 8; DB 16; Length 259;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 133 LNEKDA 140
 |||||
 Db 116 LNEKDA 123

RESULT 21
 Q7W4E3 PRELIMINARY; PRT; 265 AA.
 AC Q7W4E3 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Putative exported protein.
 GN BAI69.
 OS Bordetella bronchiseptica (Alcaligenes bronchiseptus).
 CC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 CC Alcaligenaceae; Bordetella.
 OC NCBI_TaxID=518;
 OX NCBI_TaxID=518;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RS50 / ATCC BAA-588;
 RX MEDLINE=22827954; PubMed=12910271;
 RA Parkhill J., Sebailia M., Preston A., Murphy L.D., Thomson N.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Feltham T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
 RA Leather S., Moulé S., Norberczak H., O'Neill S., Ormond D., Price C.,
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders R., Seeger K.,
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
 RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.,
 RT "Comparative analysis of the genome sequences of Bordetella pertussis,
 RT Bordetella parapertussis and Bordetella bronchiseptica.",
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL: BX640449; CA334533.1; -
 KM Complete proteome.
 SQ SEQUENCE 265 AA; 28715 MW; 4CA85474B566885B CRC64;

Query Match 2.9%; Score 8; DB 16; Length 265;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 131 LNEKDA 138
 |||||
 Db 120 LNEKDA 127

RESULT 22
 Q7W4E3 PRELIMINARY; PRT; 265 AA.
 ID Q7W4E3

AC Q7W4E3 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Putative exported protein.
 GN BAI69.
 OS Bordetella bronchiseptica (Alcaligenes bronchiseptus).
 CC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 CC Alcaligenaceae; Bordetella.
 OC NCBI_TaxID=518;
 OX NCBI_TaxID=518;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RS50 / ATCC BAA-588;
 RX MEDLINE=22827954; PubMed=12910271;
 RA Parkhill J., Sebailia M., Preston A., Murphy L.D., Thomson N.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Feltham T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
 RA Leather S., Moulé S., Norberczak H., O'Neill S., Ormond D., Price C.,
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders R., Seeger K.,
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
 RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.,
 RT "Comparative analysis of the genome sequences of Bordetella pertussis,
 RT Bordetella parapertussis and Bordetella bronchiseptica.",
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL: BX640449; CA334533.1; -
 KM Complete proteome.
 SQ SEQUENCE 265 AA; 28715 MW; 4CA85474B566885B CRC64;

Query Match 2.9%; Score 8; DB 16; Length 265;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 131 LNEKDA 138
 |||||
 Db 120 LNEKDA 127

RESULT 23
 Q7W4E3 PRELIMINARY; PRT; 265 AA.
 AC Q7W4E3 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Putative exported protein.
 GN BAI69.
 OS Bordetella bronchiseptica (Alcaligenes bronchiseptus).
 CC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 CC Alcaligenaceae; Bordetella.
 OC NCBI_TaxID=518;
 OX NCBI_TaxID=518;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RS50 / ATCC BAA-588 / NCTC 13251;
 RX MEDLINE=22827954; PubMed=12910271;
 RA Parkhill J., Sebailia M., Preston A., Murphy L.D., Thomson N.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Feltham T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
 RA Leather S., Moulé S., Norberczak H., O'Neill S., Ormond D., Price C.,
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders R., Seeger K.,
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
 RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.,
 RT "Comparative analysis of the genome sequences of Bordetella pertussis,
 RT Bordetella parapertussis and Bordetella bronchiseptica.",
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL: BX640449; CA334533.1; -
 KM Complete proteome.
 SQ SEQUENCE 265 AA; 28715 MW; 6B585474B5668844 CRC64;

Query Match 2.9%; Score 8; DB 16; Length 265;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

131 LNEKDDGA 138
 |||||
 120 LNEKDDGA 127

RESULT 24
 Q8P4S9 PRELIMINARY; PRT; 266 AA.

AC Q8P4S9;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Outer membrane protein.
 GN XCC3628.
 OS Xanthomonas campestris (pv. campestris).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 CX NCBI_TaxID=340;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 33913 / NCPPB 528;
 RC MEDLINE=22022145; PubMed=12024217;
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
 RA Camarotte G., Canavan F., Cardoso J., Chamberg F., Clapina L.P.,
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
 RA Faria U.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
 RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
 RA Setubal J.C., Kitajima J.P.;
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing
 RT host specificities.";
 RL Nature 417:459-463(2002).
 DR EMBL: AE012484; AAM42898.1;
 DR GO: GO:0016072; C:membrane; IEA.
 DR GO: GO:0004872; F:receptor activity; IEA.
 DR GO: GO:0005215; F:transporter activity; IEA.
 DR GO: GO:0006810; P:transport; IEA.
 DR InterPro: IPR004872; Lipoprotein_9.
 DR InterPro: IPR000437; Prok Lipoprot_S.
 DR InterPro: IPR000531; TonB box.
 DR Pfam: PF03180; Lipoprotein_9.1.
 DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 DR PROSITE: PS00430; TONB_DEPENDENT_REC_1; 1.
 KW Complete proteome.
 SQ SEQUENCE 266 AA; 28870 MW; 9023F0E86CD3390 CRC64;

Query Match 2.9%; Score 8; DB 16; Length 266;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

237 VARADND 244
 |||||
 227 VARADND 234

RESULT 25
 Q7V195 PRELIMINARY; PRT; 268 AA.
 AC Q7V195;
 DT 01-OCT-2003 (TREMBlrel. 25, Created)

DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Hypothetical protein yaeC.
 GN YAE_C OR H00713.
 OS Helicobacter hepaticus.
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 OC Helicobacteraceae; Helicobacter.
 CX NCBI_TaxID=32025;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 51449 / 3B1;
 RC MEDLINE=22709201; PubMed=12810954;
 RA Stuerbaum S., Josephans C., Stezenbach T., Drescher B., Brandt P.,
 RA Bell M., Droege M., Fartmann B., Fischer H.-P., Ge Z., Hoerster A.,
 RA Holland R., Klein K., Koenig J., Macko L., Mendz G.L., Nykatura G.,
 RA Schauer D.B., Shen Z., Weber J., Frosch M., Fox J.G.,
 RA "The complete genome sequence of the carcinogenic bacterium
 RT Helicobacter hepaticus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7901-7906(2003).
 DR EMBL: AE017146; AAP7310.1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 268 AA; 29314 MW; 20948E0B2918109F CRC64;

Query Match 2.9%; Score 8; DB 16; Length 268;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

175 LDIYENPK 182
 |||||
 168 LDIYENPK 175

RESULT 26
 Q8P6F0 PRELIMINARY; PRT; 269 AA.

AC Q8P6F0;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Outer membrane protein.
 GN XAC3667.
 OS Xanthomonas axonopodis (pv. citri).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 CX NCBI_TaxID=92829;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=306 / ATCC 13902 / XV 101;
 RC MEDLINE=22022145; PubMed=12024217;
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
 RA Camarotte G., Canavan F., Cardoso J., Chamberg F., Clapina L.P.,
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
 RA Faria U.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
 RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
 RA Setubal J.C., Kitajima J.P.;
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing
 RT host specificities.";
 RL Nature 417:459-463(2002).
 DR EMBL: AE012017; AAM38510.1;
 DR GO: GO:0016072; C:membrane; IEA.
 DR GO: GO:0004872; F:receptor activity; IEA.
 DR GO: GO:0005215; F:transporter activity; IEA.
 DR GO: GO:0006810; P:transport; IEA.
 DR InterPro: IPR004872; Lipoprotein_9.

DR InterPro; IPRO00437; Prok_lipoprot_S.
 DR InterPro; IPRO00531; TonB_boxC.
 DR Pfam; PF03180; Lipoprotein_9; 1.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; 1.
 KM Complete proteome.
 SQ SEQUENCE 269 AA; 29097 MW; 75F2ACE07C28806A CRC64;

Query Match 2.9%; Score 8; DB 16; Length 269;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 237 VARADNKD 244
 |||||
 DB 230 VARADNKD 237

RESULT 27
 Q87RS3 PRELIMINARY; PRT; 269 AA.
 AC Q87RS3;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Lipoprotein YaeC.
 GN VP0704.
 OS Vibrio parahaemolyticus.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrrio.
 OX NCBI_TaxID=670;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RIND 2210633 / Serotype O3:K6;
 RX MEDLINE=22508454; PubMed=12620739;
 RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
 RA Iijima Y., Najiima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
 RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
 RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
 RT distinct from that of V. cholerae.";
 RL Lancet 361:743-749 (2003).
 DR EMBL; AP05075; BACS8967.1; -
 DR InterPro; IPRO04872; Lipoprotein_9.
 DR InterPro; IPRO00437; Prok_lipoprot_S.
 DR Pfam; PF03180; Lipoprotein_9; 1.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 KM Complete proteome.
 SQ SEQUENCE 269 AA; 29070 MW; 5D458EA921862ED6 CRC64;

Query Match 2.9%; Score 8; DB 16; Length 269;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 VAKKYNL 65
 |||||
 DB 52 VAKKYNL 59

RESULT 28
 Q8CWC8 PRELIMINARY; PRT; 270 AA.
 AC Q8CWC8;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE D-methionine-binding lipoprotein meco precursor.
 GN YAEK OR C0238.
 OS Escherichia coli O6.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=217992;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=06:HI / CFT073 / ATCC 700928;

RX MEDLINE=22388233; PubMed=12471157;
 RA Welch R.A., Burland V., Plunkett G., III, Redford P., Roesch P.,
 RA Raebio D., Buckles E.L., Lion S.-R., Boutin A., Hackett J., Stroud D.,
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
 RA Mobley H.L.T., Donenberg M.S., Blattner F.R.;
 RT "Extensive mosaic structure revealed by the complete genome sequence
 RT of uropathogenic Escherichia coli.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
 DR EMBL; AE016755; AAN78730.1; -
 DR InterPro; IPRO04872; Lipoprotein_9.
 DR InterPro; IPRO00437; Prok_lipoprot_S.
 DR InterPro; IPRO04478; YaeC.
 DR Pfam; PF03180; Lipoprotein_9; 1.
 DR TIGRfams; TIGR00363; TIGR00363; 1.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 KM Lipoprotein, Complete proteome.
 SQ SEQUENCE 270 AA; 29318 MW; 2EC4133C671EAE1EC CRC64;

Query Match 2.9%; Score 8; DB 16; Length 270;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 VGNFVYP 120
 |||||
 DB 106 VGNFVYP 113

RESULT 29
 Q895C6 PRELIMINARY; PRT; 272 AA.
 AC Q895C6;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE ABC transporter substrate binding protein.
 GN CTC01352.
 OS Clostridium tetani.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Massachusetts / E88;
 RX MEDLINE=22457253; PubMed=12552129;
 RA Bregemann H., Baumer S., Fricke W.F., Wieser A., Liesegang H.,
 RA Decker I., Herzberg C., Martinez-Arias R., Merkl R., Henne A.,
 RA Gottschalk G.;
 RT "The genome sequence of Clostridium tetani, the causative agent of
 RT tetanus disease.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321 (2003).
 DR EMBL; AE015940; AAC5914.1; -
 DR InterPro; IPRO04872; Lipoprotein_9.
 DR InterPro; IPRO00437; Prok_lipoprot_S.
 DR Pfam; PF03180; Lipoprotein_9; 1.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 KM Complete proteome.
 SQ SEQUENCE 272 AA; 30259 MW; 36F2AC2CC5368C70 CRC64;

Query Match 2.9%; Score 8; DB 16; Length 272;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 KDSPIYNI 235
 |||||
 DB 224 KDSPIYNI 231

RESULT 30
 Q7Z3D8 PRELIMINARY; PRT; 277 AA.
 AC Q7Z3D8;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)

```

DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypothetical protein DKFZP686L2452 (Fragment).
GN DKFZP686L2452.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human endometrium;
RA Poustka A., Albert R., Moosmayer P., Schupp I., Wellenreuther R.,
RA Mewes H.W., Weil B., Auld C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX537968; CADD97932.1; -.
KW Hypothetical protein.
FT NOYTER
SQ SEQUENCE 277 AA; 29753 MW; 6A8691B0AEB4E7B7 CRC64;

Query Match
Best Local Similarity 100.0%; Score 8; DB 4; Length 277;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 YLEKDSQE 104
DB 158 YLEKDSQE 165

RESULT 31
ID Q9AIE4 PRELIMINARY; PRT; 281 AA.
AC Q9AIE4;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypothetical protein SPY0319.
GN SPY0319.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
RX MEDLINE=21152684; PubMed=1196296;
RA Ferretti J.J., McGhan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezesh S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren O., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RL "Complete genome sequence of an M1 strain of Streptococcus pyogenes."
RT Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
DR EMBL; AE006497; AAK3378.1; -.
DR InterPro; IPR004872; Lipoprotein 9.
DR InterPro; IPR000437; Prok lipoprot. S.
DR Pfam; PF03180; Lipoprotein 9; 1.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 281 AA; 31140 MW; 2D7E7037B0C0C270 CRC64;

Query Match
Best Local Similarity 100.0%; Score 8; DB 16; Length 281;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 KAYQTDDEV 260
DB 258 KAYQTDDEV 265

RESULT 32
ID Q8P2K7 PRELIMINARY; PRT; 281 AA.
AC Q8P2K7;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)

```

```

DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypothetical protein SPYM18_0314.
GN SPYM18_0314.
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=186103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS8232 / Serotype M18;
RX MEDLINE=21927593; PubMed=11917108;
RA Smoot J.C., Barbhan K.D., Van Gompel J.U., Smoot I.M., Chaussee M.S.,
RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks."
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
DR EMBL; AE009977; AAL97072.1; -.
DR InterPro; IPR004872; Lipoprotein 9.
DR InterPro; IPR000437; Prok lipoprot. S.
DR Pfam; PF03180; Lipoprotein 9; 1.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 281 AA; 31110 MW; 2D907187B0D8C3CB CRC64;

Query Match
Best Local Similarity 100.0%; Score 8; DB 16; Length 281;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 KAYQTDDEV 260
DB 258 KAYQTDDEV 265

RESULT 33
ID Q8K8K9 PRELIMINARY; PRT; 281 AA.
AC Q8K8K9;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Putative lipoprotein (Conserved hypothetical protein).
GN ATMB OR SPYM3_0233 OR SP31626.
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=198466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS315 / Serotype M3;
RX MEDLINE=22135806; PubMed=12122206;
RA Beres S.B., Sylva G.L., Barbhan K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Musser J.M.;
RT "Genome sequence of a serotype M3 strain of group A Streptococcus;
RT phage-encoded toxins, the high-virulence phenotype, and clone
RT emergence."
RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SSI-1 / Serotype M3;
RA Nakagawa I., Kurokawa K., Nakata M., Tomiyasu Y., Yamashita A.,
RA Yamazaki K., Okahashi N., Kawabata S., Yasunaga T., Hattori M.,
RA Hayashi H., Hamada S.;
RT "The genome of invasive Streptococcus pyogenes; a comparative analysis
RT of S. pyogenes SSI-1, SF370 and MGAS8232."
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE014141; AAM78840.1; -.
DR EMBL; AP005146; BAC64721.1; -.
DR InterPro; IPR004872; Lipoprotein 9.

```


DR InterPro: IPR000437; Prok_Lipoprot_S.
 DR Pfam: PF03180; Lipoprotein_7.1
 DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 KM Hypothetical protein; Lipoprotein; Complete proteome.
 SQ SEQUENCE 281 AA; 3111 MW; 7D907B87B76D1510 CRC64;

Query Match 2.9%; Score 8; DB 16; Length 281;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 KAYOTDEV 260
 DB 258 KAYOTDEV 265

RESULT 34
 Q9P876 PRELIMINARY; PRT; 305 AA.

AC Q9P876; 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Mitochondrial ADP/ATP carrier isoform 1.
 GN ANC.
 OS Pichia jadinii (Yeast) (Candida utilis).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Pichia.
 OX NCBI_TaxID=4903;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 9950;
 RA David C., Terequet V., Gonzalez C., Languin G.M.;
 RT "Several genes encode the mitochondrial ADP/ATP carrier in the petite
 negative yeast Candida utilis."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 CC -1 SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
 DR EMBL; AJ277098; CAB88027.1; -
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005743; C:mitochondrial inner membrane; IEA.
 DR GO; GO:0005488; F:binding; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro: IPR001993; Mitoch carrier.
 DR InterPro: IPR002067; Mit_carr; 3.
 DR Pfam; PF00153; mto_carr; 3.
 DR PRINTS; PR00926; MITOCARRIER.
 DR PROSITE; PS00215; MITOCH_CARRIER; 2.
 KM Membrane; Transmembrane; Transport.
 SQ SEQUENCE 305 AA; 33202 MW; 9E819E8A0D1AA601 CRC64;

Query Match 2.9%; Score 8; DB 3; Length 305;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 AAISKTA 36
 DB 21 AAISKTA 28

RESULT 35
 Q9P875 PRELIMINARY; PRT; 305 AA.

AC Q9P875; 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Mitochondrial ADP/ATP carrier isoform 2.
 GN ANC.
 OS Pichia jadinii (Yeast) (Candida utilis).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Pichia.
 OX NCBI_TaxID=4903;
 RN (1)
 RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 9950;
 RA David C., Terequet V., Gonzalez C., Languin G.M.;
 RT "Several genes encode the mitochondrial ADP/ATP carrier in the petite
 negative yeast Candida utilis."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 CC -1 SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
 DR EMBL; AJ277099; CAB88028.1; -
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005743; C:mitochondrial inner membrane; IEA.
 DR GO; GO:0005488; F:binding; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro: IPR001993; Mitoch carrier.
 DR InterPro: IPR002067; Mit_carr; 3.
 DR Pfam; PF00153; mto_carr; 3.
 DR PRINTS; PR00926; MITOCARRIER.
 DR PROSITE; PS00215; MITOCH_CARRIER; 2.
 KM Membrane; Transmembrane; Transport.
 SQ SEQUENCE 305 AA; 33188 MW; 72F0596160AC7D3C CRC64;

Query Match 2.9%; Score 8; DB 3; Length 305;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 AAISKTA 36
 DB 21 AAISKTA 28

RESULT 36
 Q8ZSM1 PRELIMINARY; PRT; 347 AA.

AC Q8ZSM1; 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Hypothetical protein PAE3560.
 GN PAE3560.
 OS Pyrobaculum aerophilum.
 OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
 CC Thermoproteales; Pyrobaculum.
 OX NCBI_TaxID=13773;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
 RX MEDLINE=21664397; PubMed=11792869;
 RA Fitz-Gibbon S.T., Ladner H., Kim U.-U., Stetter K.O., Simon M.I.,
 RA Miller J.H.;
 RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
 aerophilum."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
 DR EMBL; AE009937; AAL65002.1; -
 KM Hypothetical protein; Complete proteome.
 SQ SEQUENCE 347 AA; 38630 MW; 5BCF4154F54420D0 CRC64;

Query Match 2.9%; Score 8; DB 17; Length 347;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 198 IDVDPLAV 205
 DB 294 IDVDPLAV 301

RESULT 37
 Q8A673 PRELIMINARY; PRT; 351 AA.

AC Q8A673; 01-JUN-2003 (TREMBlrel. 24, Created)
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Putative phosphohydrolases.
 GN BT2013.
 OS Bacteroides thetaiotaomicron.

OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
 OC Bacteroidaceae; Bacteroides.
 OX NCBI_TaxID=818;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VPI-5482 / ATCC 29148;
 RX MEDLINE=22550858; PubMed=12663928;
 RA Xu Y., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,
 RA Chang H.C., Hooper L.V., Gordon J.I.;
 RT "A genomic view of the human-bacteroides thetaiotaomicron symbiosis";
 RL Science 299:2074-2076(2003).
 DR EMBL: AE016934; AAC77120.1;
 DR GO: GO:0016787; F:hydrolase activity; IEA.
 DR InterPro: IPR004643; M-peptidase.
 DR InterPro: IPR000437; F:ox_lipoprot_S.
 DR Pfam: PF00149; Metallophos; 1.
 DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 KW Hydrolase; Complete proteome.
 SQ SEQUENCE 351 AA; 39029 MW; 9FB50645491BB86 CRC64;

Query Match 2.9%; Score 8; DB 16; Length 351;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 104 EKGJNNLV 111
 |||||
 DB 142 EKGJNNLV 149

RESULT 38
 OQ3AL8 PRELIMINARY; PRT; 374 AA.
 ID OQ3AL8;
 AC OQ3AL8;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE R-benylactate dehydratase small subunit.
 GN PLDC.
 OS Clostridium sporogenes.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1509;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 3584;
 RA Dichter S., Pierik A.J., Buckel W.;
 RT "Molecular characterization of (R)-phenyllactate dehydratase and its
 RT activating component A from Clostridium sporogenes";
 RL Submitted (SSP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF420489; AAL18811.1;
 SQ SEQUENCE 374 AA; 43148 MW; 84297814477B60B0 CRC64;

Query Match 2.9%; Score 8; DB 2; Length 374;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 176 DIVENPKK 183
 |||||
 DB 13 DIVENPKK 20

RESULT 39
 OQHYEO PRELIMINARY; PRT; 446 AA.
 ID OQHYEO;
 AC OQHYEO;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Probable ATP-dependent RNA helicase.
 GN PA3466.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.

OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=29437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mitzoguchi S.D., Watterer P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Gabor R.L., Golty L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brady L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Slier M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RT opportunistic pathogen";
 RL Nature 406:959-964(2000).
 CC -1- SIMILARITY: TO DEAD/DEAH BOX HELICASE FAMILY.
 CC -1- SIMILARITY: TO HELICASE C-TERMINAL DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE "DEAD" BOX FAMILY HELICASES.
 DR EMBL: AE004767; AAC6854.1;
 DR PIR: A83213; A83213.
 DR HSP, Q58083; HVH.
 DR GO: GO:0005622; C:intracellular; IEA.
 DR GO: GO:0005524; F:ATP binding; IEA.
 DR GO: GO:0008026; F:ATP dependent helicase activity; IEA.
 DR GO: GO:0003723; F:RNA binding; IEA.
 DR GO: GO:0003700; F:transcription factor activity; IEA.
 DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR006628; DEAD_box.
 DR InterPro: IPR001650; Helicase_C.
 DR InterPro: IPR000005; HTHfac.
 DR Pfam: PF00270; DEAD; 1.
 DR Pfam: PF00271; Helicase_C; 1.
 DR SMART: SM00487; DEXDC; 1.
 DR SMART: SM00490; HELIC; 1.
 DR PROSITE: PS00039; DEAD_ATP_HELICASE; 1.
 DR PROSITE: PS00041; HTH_ARAC_FAMILY; 1.
 KW ATP-binding; Helicase; RNA-binding; Complete proteome.
 SQ SEQUENCE 446 AA; 48989 MW; 816D6CD046851C6 CRC64;

Query Match 2.9%; Score 8; DB 16; Length 446;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 149 LARALLIL 156
 |||||
 DB 76 LARALLIL 83

RESULT 40
 OQ9R04 PRELIMINARY; PRT; 512 AA.
 ID OQ9R04;
 AC OQ9R04;
 DT 01-JUN-2003 (TREMBlrel. 24, Created)
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE NAD-dependent succinate aldehyde dehydrogenases.
 GN ARXK OR BL2968.
 OS Bradyrhizobium japonicum.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Bradyrhizobiaceae; Bradyrhizobium.
 OX NCBI_TaxID=375;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=USDA 110;
 RX MEDLINE=22484998; PubMed=12597275;
 RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiino T.,
 RA Sasamoto S., Watanabe A., Idesawa K., Iizuchi M., Kawasumi K.,
 RA Kohara M., Matsumoto M., Shimo S., Tsuruoka H., Wada T., Yamada M.,
 RA Tabata S.;
 RT "Complete genome sequence of nitrogen-fixing symbiotic bacterium
 RT Bradyrhizobium japonicum USDA110";
 RL DNA Res. 9:189-197(2002).
 DR EMBL: AP005945; BAC48233.1;

DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:008152; P:metabolism; IEA.
DR InterPro; IPR02086; Aldehyde_dehydr.
DR Pfam; PF0171; aldedh; 1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; 1.
KW Complete proteome.
SQ SEQUENCE 512 AA; 55015 MW; 120B682457B83COA CRC64;

Query Match

Best Local Similarity 2.9%; Score 8; DB 16; Length 512;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 135 KDQATIAV 142
Db 367 KDQATIAV 374

Search completed: June 16, 2004, 11:17:15
Job time : 65 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 16, 2004, 11:08:47 / Search time 61 seconds
(without alignments)
1278.412 Million cell updates/sec

Title: US-10-018-672-2

Sequence: 1 MNRGKINGICALASGIALAG.....TDEVEAEAKKQKPDGVIKGM 276

Scoring table: OLIGO
Gapop 60.0, Gapext 60.0

Searched: 1586107 seqs, 282547505 residues

Word size: 0

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database: A_Geneseq_29Jan04:**

1: Geneseq1980s:**
2: Geneseq1990s:**
3: Geneseq2000s:**
4: Geneseq2001s:**
5: Geneseq2002s:**
6: Geneseq2003as:**
7: Geneseq2003bs:**
8: Geneseq2004s:**

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	276	100.0	276	4	AA820104
2	243	88.0	276	5	AA017582
3	243	88.0	276	6	ABU35462
4	19	6.9	272	5	AA017813
5	19	6.9	273	4	AAU35474
6	19	6.9	273	5	AB876073
7	19	6.9	273	5	AB876073
8	19	6.9	273	5	AB876073
9	16	5.8	270	5	AB876094
10	16	5.8	271	6	ABU40467
11	14	5.1	20	5	AB876095
12	14	4.7	270	6	ABU24914
13	14	4.3	276	6	ABU39463
14	14	4.0	67	4	AA870412
15	11	4.0	271	6	ABU27902
16	11	3.6	276	6	ABU16868
17	11	3.6	276	6	ABU16868
18	10	3.6	125	6	AB853515
19	9	3.3	241	6	ABU16864
20	9	3.3	256	6	ABU33203
21	9	3.3	262	6	ABU23447
22	9	3.3	270	6	ABU27880
23	9	3.3	272	6	ABM67576
24	9	3.3	273	5	AB848120
25	9	3.3	273	5	AB848120

26	9	3.3	273	6	ABU32634
27	9	3.3	275	6	ABU49250
28	9	3.3	291	7	ADG97607
29	9	3.3	296	6	ADA35637
30	8	2.9	96	4	AAU18017
31	8	2.9	96	4	AB810497
32	8	2.9	96	4	AB876084
33	8	2.9	96	7	AD831641
34	8	2.9	152	4	AA005643
35	8	2.9	240	4	AAU36063
36	8	2.9	263	5	ABP26478
37	8	2.9	263	5	ABU25052
38	8	2.9	265	6	ABU23277
39	8	2.9	269	7	ADG6816
40	8	2.9	270	6	AAU21415
41	8	2.9	271	4	AAU34457
42	8	2.9	271	4	ABU31676
43	8	2.9	271	6	ABU15367
44	8	2.9	281	6	ABU46465
45	8	2.9	282	6	ABU29987
46	8	2.9	295	6	ABU19780
47	8	2.9	379	4	AB810313
48	8	2.9	379	5	ABP66900
49	8	2.9	739	5	AA668310
50	8	2.9	744	5	AA668310
51	8	2.9	758	4	AB858337
52	8	2.9	766	5	AA668312
53	8	2.9	771	5	AA668311
54	8	2.9	838	5	AA668302
55	8	2.9	843	5	AA668301
56	8	2.9	865	5	AA668304
57	8	2.9	870	5	AA668303
58	8	2.5	9	5	AB876087
59	7	2.5	9	5	AB876093
60	7	2.5	9	5	AB876086
61	7	2.5	9	5	AB876092
62	7	2.5	73	2	AAW20141
63	7	2.5	78	2	AAW38625
64	7	2.5	92	4	AAW25520
65	7	2.5	98	5	ABP07971
66	7	2.5	105	4	AA881880
67	7	2.5	114	7	ADCO0063
68	7	2.5	114	7	AD899989
69	7	2.5	114	7	ADCO0470
70	7	2.5	114	7	ADCO0739
71	7	2.5	114	7	ADCO0848
72	7	2.5	114	7	ADCO1030
73	7	2.5	114	7	ADCO0319
74	7	2.5	116	5	ABP38486
75	7	2.5	117	5	ABP30711
76	7	2.5	131	3	AA857230
77	7	2.5	132	3	AA844014
78	7	2.5	151	5	ABP30811
79	7	2.5	153	2	AAW20166
80	8	2.5	159	4	AA803580
81	7	2.5	161	2	AA899625
82	7	2.5	172	5	ABP28372
83	7	2.5	173	4	AAU29665
84	7	2.5	174	7	AD868159
85	7	2.5	178	7	AD868153
86	7	2.5	186	4	AB869789
87	7	2.5	194	4	AAU14092
88	7	2.5	194	4	AB833037
89	7	2.5	194	4	AAU26499
90	7	2.5	194	4	AB827866
91	7	2.5	194	4	AB818510
92	7	2.5	194	4	AAW66222
93	7	2.5	194	4	AAW38336
94	7	2.5	194	4	AAW47890
95	7	2.5	194	4	AAW47890
96	7	2.5	194	5	ABG35872
97	7	2.5	197	4	ABG16733
98	7	2.5	199	2	AA860986

ABU32634	Protein e
ABU49250	Protein e
ADG97607	E. faeciu
ADA35637	Acinetoba
AAU18017	Human imm
AB810497	Human CDN
AB876084	Human pol
AD831641	Human nov
AA005643	Human pol
AAU36063	Klebsiell
ABP26478	Streptoco
ABU25052	Protein e
ABU23277	Protein e
ADG6816	E. faeciu
AAU21415	Protein e
AAU34457	E. coli c
ABU31676	Protein e
ABU15367	Protein e
ABU46465	Protein e
ABU29987	Protein e
ABU19780	Protein e
AB810313	Human CDN
ABP66900	Human pol
AA668310	Human sem
AA668310	Human sem
AA668309	Human sem
AB858337	Drosophil
AA668312	Human sem
AA668311	Human sem
AA668302	Human sem
AA668301	Human sem
AA668304	Human sem
AA668303	Human sem
AB876087	Haemophil
AB876093	Haemophil
AB876086	Haemophil
AB876092	Haemophil
AAW20141	H. pylori
AAW38625	Streptoco
AAW25520	Human pro
ABP07971	Human ORF
AA881880	S. epider
ADCO0063	Enterohae
AD899989	Enterohae
ADCO0470	Enterohae
ADCO0739	Enterohae
ADCO0848	Enterohae
ADCO1030	Enterohae
ADCO0319	Enterohae
ABP38486	Staphyloc
ABP30711	Streptoco
AA857230	Human pro
AA844014	Human can
ABP30811	Streptoco
AAW20166	Helicobac
AA803580	Human pro
AA899625	P. aerugi
ABP28372	Streptoco
AAU29665	Novel hum
AD868159	Human GPC
AD868153	Human GPC
AB869789	Drosophil
AAU14092	Peptide #
AB833037	Peptide #
AAU26499	Peptide #
AB827866	Human pep
AB818510	Protein #
AAW66222	Human bon
AAW38336	Human bra
AAW47890	Human liv
AAW47890	Peptide #
ABG35872	Human pep
ABG16733	Novel hum
AA860986	A 22-26 x

99	2.5	214	4	ABG00371	Novel hum	172	7	2.5	350	1	ABP2053	ApB2053	Outer mem
100	2.5	220	6	ABU41510	Protein e	173	7	2.5	350	4	ABP5179	ABP5179	Protein a
101	2.5	220	7	ADC39190	Novel hum	174	7	2.5	352	4	AAU34460	AAU34460	E. coli c
102	2.5	222	4	ABG27431	Novel hum	175	7	2.5	352	6	ABM67622	ABM67622	Phototrab
103	2.5	222	5	ABP74106	Human TRI	176	7	2.5	355	7	ADC86179	ADC86179	Human GPC
104	2.5	226	2	ABP96627	P. aerugi	177	7	2.5	356	4	AAE03581	AAE03581	Human pro
105	2.5	228	5	ABP43305	Human ova	178	7	2.5	411	7	ADC46587	ADC46587	Chlamydia
106	2.5	241	6	ABD07136	Alliolooc	179	7	2.5	419	2	AAW25465	AAW25465	Mycobacte
107	2.5	244	4	ABE80307	Corynebac	180	7	2.5	419	6	AAW81240	AAW81240	Mycobacte
108	2.5	244	5	ABG80339	C. glutam	181	7	2.5	419	6	ABU37016	ABU37016	Protein e
109	2.5	247	6	ABG26453	Protein e	182	7	2.5	424	4	AAW90552	AAW90552	Protein e
110	2.5	256	6	ABU39703	Protein e	183	7	2.5	424	7	AAW90552	AAW90552	C. glutam
111	2.5	260	4	AAU36517	Pseudomon	184	7	2.5	429	6	ABR41715	ABR41715	Human pro
112	2.5	260	6	ABU38920	Protein e	185	7	2.5	438	4	ABE03582	ABE03582	Human pro
113	2.5	261	6	ABU43388	Protein e	186	7	2.5	460	4	AAW90469	AAW90469	C. glutam
114	2.5	262	6	ABU49395	Protein e	187	7	2.5	460	4	AAW90469	AAW90469	C. glutam
115	2.5	262	6	ABU18705	Protein e	188	7	2.5	468	4	AAW90574	AAW90574	Human pro
116	2.5	263	4	ABG27056	Novel hum	189	7	2.5	475	2	AAW31046	AAW31046	Kat DIB d
117	2.5	265	6	ABU39208	Protein e	190	7	2.5	477	2	AAW21082	AAW21082	Dopamine
118	2.5	268	2	AAW90682	Rat calde	191	7	2.5	477	2	AAW79381	AAW79381	Dopamine
119	2.5	268	6	ABU48077	Protein e	192	7	2.5	477	2	AAW40802	AAW40802	Human DS
120	2.5	268	6	ABU27644	Protein e	193	7	2.5	477	2	ABW56350	ABW56350	Non-endog
121	2.5	268	6	ABU21332	Protein e	194	7	2.5	477	6	ABP81808	ABP81808	Human dop
122	2.5	268	6	ABU47067	Protein e	195	7	2.5	479	2	AAW25546	AAW25546	Truncated
123	2.5	268	6	ABU14874	Protein e	196	7	2.5	480	5	ABG78365	ABG78365	Human pro
124	2.5	268	6	ABU45650	Protein e	197	7	2.5	496	6	AAW81966	AAW81966	S. epider
125	2.5	269	6	ABU30644	Protein e	198	7	2.5	496	6	ABU45063	ABU45063	Protein e
126	2.5	269	6	ABU41046	Protein e	199	7	2.5	505	4	ABG15313	ABG15313	Novel hum
127	2.5	269	6	ABU31515	Protein e	200	7	2.5	506	5	ABP39538	ABP39538	Staphyloc
128	2.5	270	6	ABM69136	Phototrab	201	7	2.5	507	2	AAE03576	AAE03576	Human pro
129	2.5	270	6	ABU17607	Protein e	202	7	2.5	519	5	ABP65572	ABP65572	Human pro
130	2.5	271	2	AAU10962	H. pylori	203	7	2.5	527	4	ABW61819	ABW61819	Drosophila
131	2.5	271	2	AAU10998	H. pylori	204	7	2.5	556	5	ABP73756	ABP73756	Candida a
132	2.5	271	2	AAW89937	Actinogen	205	7	2.5	575	5	AAU17811	AAU17811	H. influenza
133	2.5	271	2	AAW89833	Protein e	206	7	2.5	575	5	AAE03577	AAE03577	Human pro
134	2.5	271	2	AAW90023	Expressed	207	7	2.5	611	4	AAW96570	AAW96570	Putative
135	2.5	271	4	AAU35867	Helicobac	208	7	2.5	633	6	ABU26147	ABU26147	Protein e
136	2.5	271	4	AAU35867	Helicobac	209	7	2.5	652	2	AAW80610	AAW80610	S. pneumo
137	2.5	271	6	ABU50250	Protein e	210	7	2.5	652	4	AAU38080	AAU38080	Streptococ
138	2.5	271	6	ABU50250	Protein e	211	7	2.5	652	4	AAU37724	AAU37724	Streptococ
139	2.5	271	6	ABU43419	Protein e	212	7	2.5	652	6	ABU01536	ABU01536	S. pneumo
140	2.5	271	6	ABU31110	Protein e	213	7	2.5	652	6	ABU46049	ABU46049	Drosophila
141	2.5	272	6	ABU24666	Protein e	214	7	2.5	668	4	ABW61757	ABW61757	Protein e
142	2.5	274	6	ABU50032	Protein e	215	7	2.5	670	4	ABW71706	ABW71706	Drosophila
143	2.5	276	5	ABW49303	Protein e	216	7	2.5	670	6	ABU19181	ABU19181	Protein e
144	2.5	276	6	ABU32999	Listeria	217	7	2.5	675	4	ABW59812	ABW59812	Drosophila
145	2.5	280	2	AAW20909	H. pylori	218	7	2.5	689	4	ABW68912	ABW68912	Drosophila
146	2.5	286	2	AAW85996	S. pneumo	219	7	2.5	691	7	ABW00766	ABW00766	Helicobac
147	2.5	286	5	ABW53624	Lactococc	220	7	2.5	691	2	AAW73027	AAW73027	Helicobac
148	2.5	286	5	ABW53625	Lactococc	221	7	2.5	691	2	AAW73032	AAW73032	Helicobac
149	2.5	293	3	ADC39192	Novel hum	222	7	2.5	691	7	ABW00781	ABW00781	Helicobac
150	2.5	298	3	AAV71032	Human ade	223	7	2.5	692	4	AAW33225	AAW33225	C. glutam
151	2.5	298	3	AAV71033	Human ade	224	7	2.5	717	1	AAW33225	AAW33225	C. glutam
152	2.5	298	3	AAU01199	Human ade	225	7	2.5	757	4	AAW33225	AAW33225	C. glutam
153	2.5	298	4	AAU01200	Human ade	226	7	2.5	757	4	AAW33225	AAW33225	C. glutam
154	2.5	298	4	AAU01200	Human ade	227	7	2.5	766	4	AAW33225	AAW33225	C. glutam
155	2.5	298	5	AAU01200	Human ade	228	7	2.5	852	6	ABP79237	ABP79237	N. gonorr
156	2.5	298	5	AAU01200	Human ade	229	7	2.5	852	6	ABU21216	ABU21216	Human pro
157	2.5	298	5	AAU01200	Human ade	230	7	2.5	852	6	ABU37147	ABU37147	Protein e
158	2.5	302	5	ABP38326	Staphyloc	231	7	2.5	852	6	ABU37147	ABU37147	Protein e
159	2.5	307	5	ABW53940	Lactococc	232	7	2.5	858	4	AAW27282	AAW27282	Human ADA
160	2.5	309	3	AAE18898	A. maize c	233	7	2.5	871	3	AAW77289	AAW77289	C. glutam
161	2.5	310	4	AAE03575	Human pro	234	7	2.5	871	3	AAW77289	AAW77289	C. glutam
162	2.5	311	6	ABU06082	N. mening	235	7	2.5	875	5	ABW48886	ABW48886	KSG-ATG 1
163	2.5	311	6	ABU37551	Protein e	236	7	2.5	875	5	ABW48886	ABW48886	Listeria
164	2.5	316	6	ABP7807	N. gonorr	237	7	2.5	908	6	ABU32754	ABU32754	Protein e
165	2.5	316	6	ABP7807	N. gonorr	238	7	2.5	908	6	ABU32754	ABU32754	Protein e
166	2.5	317	6	AAE03579	Human pro	239	7	2.5	959	4	AAW86947	AAW86947	Human met
167	2.5	323	4	AAW41427	Human pol	240	7	2.5	1036	6	ABU49046	ABU49046	Protein e
168	2.5	325	4	ABG15423	Novel hum	241	7	2.5	1120	6	ABG76502	ABG76502	DNA encod
169	2.5	336	4	ABG21669	Novel hum	242	7	2.5	1128	6	ABU41214	ABU41214	Protein e
170	2.5	337	4	ABW05904	Microsom	243	7	2.5	1175	4	AAW5065	AAW5065	Drosophila
171	2.5	337	6	ABP99233	Ochrosomy	244	7	2.5	1175	5	ABP73177	ABP73177	Candida a

245	7	2.5	1178	6	ABU24510	Protein e	318	6	2.2	34	2	AAR40887	Aar40887 Sequence
246	7	2.5	1233	4	AAG30558	C. glutam	319	6	2.2	34	2	AAR40888	Aar40888 Sequence
247	7	2.5	1233	7	ADD13972	glut	320	6	2.2	34	2	AAR40886	Aar40886 Sequence
248	7	2.5	1239	5	AAO21545	Human pol	321	6	2.2	34	2	AAW35238	AAW35238 Human 1eu
249	7	2.5	1241	3	AAAB42626	Human ORF	322	6	2.2	34	2	AAW35240	AAW35240 Human 1eu
250	7	2.5	1263	2	AAW22052	DNA polym	323	6	2.2	34	2	AAW35241	AAW35241 Human 1eu
251	7	2.5	1482	3	AAW84727	Amino aci	324	6	2.2	34	2	AAW35237	AAW35237 Human 1eu
252	7	2.5	1488	3	AAW84725	Amino aci	325	6	2.2	34	2	AAW35239	AAW35239 Human 1eu
253	7	2.5	1509	3	AAW84728	Amino aci	326	6	2.2	35	6	ABP68391	ABP68391 Human col
254	7	2.5	1517	3	AAW84726	Amino aci	327	6	2.2	37	4	ABG28355	ABG28355 Novel hum
255	7	2.5	1615	4	AAU03536	Human pro	328	6	2.2	40	3	AAW12869	AAW12869 Degenrat
256	7	2.5	1972	2	AAW81171	Human BAZ	329	6	2.2	42	2	AAW41874	AAW41874 Peptide u
257	7	2.5	1972	6	ABR64241	Angiogene	330	6	2.2	42	5	ABW05820	ABW05820 Amb a I M
258	7	2.5	2154	6	ABU38771	Protein e	331	6	2.2	42	5	ABW05892	ABW05892 Amb a I M
259	7	2.5	3241	6	ABU22000	Protein e	332	6	2.2	43	2	AAW54448	AAW54448 Amb a I.1
260	7	2.5	4688	6	ABU48941	Protein e	333	6	2.2	43	2	AAW02503	AAW02503 Peptide c
261	7	2.5	6310	5	ABU39865	Protein e	334	6	2.2	43	4	AAW83618	AAW83618 Human imm
262	7	2.2	7	5	ABW76084	Haemophil	335	6	2.2	43	6	ABW83844	ABW83844 Drosophil
263	6	2.2	8	4	AAU04099	Human CLA	336	6	2.2	43	6	ABW83843	ABW83843 Drosophil
264	6	2.2	9	5	ABW76097	Haemophil	337	6	2.2	44	2	AAW54444	AAW54444 Amb a I.1
265	6	2.2	9	5	ABW76091	F. protein	338	6	2.2	44	2	AAW02499	AAW02499 Peptide c
266	6	2.2	10	5	ABW98527	F. protein	339	6	2.2	45	2	AAW02499	AAW02499 Peptide c
267	6	2.2	10	5	ABW98528	F. protein	340	6	2.2	45	2	AAW41873	AAW41873 Peptide u
268	6	2.2	10	5	ABW98531	F. protein	341	6	2.2	45	3	ABW25733	ABW25733 Human sec
269	6	2.2	10	5	ABW98530	F. protein	342	6	2.2	45	3	ABW05819	ABW05819 Amb a I A
270	6	2.2	10	5	ABW98529	F. protein	343	6	2.2	45	5	ABW05891	ABW05891 Amb a I A
271	6	2.2	11	4	AAW21957	Glycopept	344	6	2.2	48	3	AAW04829	AAW04829 Arabidops
272	6	2.2	11	4	AAU04100	Human ADP	345	6	2.2	49	4	ABG24210	ABG24210 Novel hum
273	6	2.2	11	5	ABW67601	Human CLA	346	6	2.2	50	4	AAW56678	AAW56678 Human nov
274	6	2.2	11	6	ADA23719	Alzheimer	347	6	2.2	50	5	ABW02306	ABW02306 Human ORF
275	6	2.2	12	5	ABW76098	Haemophil	348	6	2.2	50	6	ABW53197	ABW53197 Propionib
276	6	2.2	12	5	ABW76099	Haemophil	349	6	2.2	50	7	ABW01216	ABW01216 Enterococ
277	6	2.2	13	4	ADD23878	Breast ca	350	6	2.2	52	4	AAW91855	AAW91855 Human dig
278	6	2.2	13	4	AAW93371	Vaccine r	351	6	2.2	52	4	AAU22484	AAU22484 Novel hum
279	6	2.2	14	4	AAW72007	NAM mltog	352	6	2.2	52	7	ADB32324	ADB32324 Human nov
280	6	2.2	14	4	AAW72007	Internal	353	6	2.2	54	2	AAW77560	AAW77560 Staphyloc
281	6	2.2	15	5	AAW37897	Formylpep	354	6	2.2	54	2	AAW78358	AAW78358 M. arthri
282	6	2.2	16	5	ABW07564	Linker IT	355	6	2.2	54	4	AAW63866	AAW63866 Propionib
283	6	2.2	17	1	AAW82373	Immunosup	356	6	2.2	54	6	ABW59985	ABW59985 Propionib
284	6	2.2	17	1	AAW82370	Immunosup	357	6	2.2	57	2	AAW5826	AAW5826 S. pneumo
285	6	2.2	17	1	AAW82372	Immunosup	358	6	2.2	57	6	AAW14831	AAW14831 Human int
286	6	2.2	17	2	AAW07352	RAE26.15	359	6	2.2	59	4	AAW14833	AAW14833 Human int
287	6	2.2	17	2	AAW41907	Peptide u	360	6	2.2	59	4	AAW46397	AAW46397 Propionib
288	6	2.2	17	2	AAW58735	Allergen	361	6	2.2	59	4	AAW52789	AAW52789 Propionib
289	6	2.2	17	5	ABW05855	Ragweed p	362	6	2.2	59	6	ABW42816	ABW42816 Propionib
290	6	2.2	20	2	AAW805927	S. aureus	363	6	2.2	61	2	AAW38547	AAW38547 Streptoco
291	6	2.2	20	2	AAW85291	Helper T-	365	6	2.2	61	4	AAW89456	AAW89456 Human imm
292	6	2.2	21	1	AAW20042	Haapten re	366	6	2.2	63	4	ABW66776	ABW66776 Drosophil
293	6	2.2	21	1	AAW30017	Human lym	367	6	2.2	64	4	AAW45989	AAW45989 Propionib
294	6	2.2	21	1	AAW54428	AMB 3-4.1	368	6	2.2	64	6	AAW42508	AAW42508 Propionib
295	6	2.2	22	2	AAW54377	RAE 66.1-	369	6	2.2	66	7	AAW02898	AAW02898 Fragment
296	6	2.2	22	2	AAW02461	RAE 66.1-	370	6	2.2	66	7	AAW02898	AAW02898 Fragment
297	6	2.2	24	2	AAW02461	RAE 66.1-	371	6	2.2	67	5	AAW98428	AAW98428 Cadium-re
298	6	2.2	24	2	AAW02411	AMB 3-4.1	372	6	2.2	69	3	AAW39334	AAW39334 Arabidops
299	6	2.2	24	2	AAW02411	AMB 3-4.1	373	6	2.2	71	3	AAW61078	AAW61078 Arabidops
300	6	2.2	25	2	AAW02494	AMB 3-4.1	374	6	2.2	72	2	AAW08064	AAW08064 Gas vesic
301	6	2.2	25	2	AAW02494	AMB 3-4.1	375	6	2.2	73	5	ABW09239	ABW09239 Human ORF
302	6	2.2	25	2	AAW02460	RAE 66.1-	376	6	2.2	75	4	AAW08473	AAW08473 Human imm
303	6	2.2	27	2	AAW54427	AMB 3-5.1	377	6	2.2	75	4	AAW06742	AAW06742 Human pol
304	6	2.2	27	2	AAW02410	AMB 3-5.1	378	6	2.2	76	4	ABW02551	ABW02551 Novel hum
305	6	2.2	28	2	AAW69259	DMS prote	379	6	2.2	77	3	AAW11804	AAW11804 Arabidops
306	6	2.2	28	2	AAW54375	RAE 66.1	380	6	2.2	78	4	AAW22981	AAW22981 Novel hum
307	6	2.2	28	7	ADW48198	Putative	381	6	2.2	78	5	ABW03410	ABW03410 Human ORF
308	6	2.2	29	4	ABW24695	Novel hum	382	6	2.2	79	2	AAW32889	AAW32889 Human bra
309	6	2.2	29	6	ADW04800	IGF-IR bl	383	6	2.2	79	4	AAW32889	AAW32889 Propionib
310	6	2.2	30	6	AAW41878	Peptide u	384	6	2.2	79	6	ABW35808	ABW35808 Propionib
311	6	2.2	30	5	ABW29573	Streptoco	385	6	2.2	80	6	AAW07976	AAW07976 Allicococ
312	6	2.2	30	5	ABW05824	AGE MAJOR	386	6	2.2	81	4	AAW03486	AAW03486 Human pol
313	6	2.2	30	5	ABW05896	AGE MAJOR	387	6	2.2	81	6	ABW24670	ABW24670 Protein e
314	6	2.2	34	2	AAW40885	Sequence	388	6	2.2	82	4	AAW52720	AAW52720 Propionib
315	6	2.2	34	2	AAW40881	N-termina	389	6	2.2	82	5	ABW09448	ABW09448 Human ORF
316	6	2.2	34	2	AAW40881	Sequence	390	6	2.2	82	6	ABW49239	ABW49239 Propionib
317	6	2.2	34	2	AAW40884	Sequence	390	6	2.2	82	6	ABW49239	ABW49239 Propionib

391	6	ABU07664	Human hap	464	6	2.2	115	6	ABU01743	Abu01743 S. pneumo
392	6	ADCC00634	Enterocae	465	6	2.2	115	6	ABR83559	ABR83559 Tagged To
393	6	AAU46356	Propionib	466	6	2.2	116	6	AEU45107	AEU45107 Protein e
394	6	AAU47491	Propionib	467	6	2.2	116	7	ADC89438	ADC89438 Ribosomal
395	6	ABP09895	Human ORF	468	6	2.2	117	2	AAK52053	AAK52053 Heavy cha
396	6	ABBA49563	Listeria	469	6	2.2	117	5	ABP35225	ABP35225 Human ORF
397	6	ABMA42875	Propionib	470	6	2.2	117	7	ADC88479	ADC88479 Ribosomal
398	6	ABMA4010	Propionib	471	6	2.2	118	7	AAU08460	AAU08460 Rat Ptx2
399	6	ABU32482	Protein e	472	6	2.2	118	7	ADC87894	ADC87894 Ribosomal
400	6	AAU21455	Human nov	473	6	2.2	119	1	AAU20008	AAU20008 Hybrid hu
401	6	AAU4827	Arabisdops	474	6	2.2	119	4	AAU67797	AAU67797 Propionib
402	6	AAU16177	Human nov	475	6	2.2	119	5	AAE19376	AAE19376 Human rep
403	6	ABUS5246	Human nov	476	6	2.2	119	5	ABE78137	ABE78137 Consensus
404	6	ABP42513	Human ova	477	6	2.2	119	5	ABR07575	ABR07575 Reprolysi
405	6	ADBI0041	Human ova	478	6	2.2	119	5	ABM64316	ABM64316 Propionib
406	6	AAU57673	Arabisdops	479	6	2.2	119	6	ABK36698	ABK36698 HMM deriv
407	6	AAU16592	Human nov	480	6	2.2	119	6	AEU08384	AEU08384 Consensus
408	6	AAU44672	Propionib	481	6	2.2	120	2	AAW08366	AAW08366 RSV F pro
409	6	ABBI1429	Human ner	482	6	2.2	120	2	ABG12411	ABG12411 Novel hum
410	6	ABM41191	Propionib	483	6	2.2	120	5	ABE04303	ABE04303 Recombina
411	6	ABUS5661	Human nov	484	6	2.2	121	4	AAO00446	AAO00446 Human pol
412	6	AAU39606	Arabisdops	485	6	2.2	121	5	ABP09591	ABP09591 Human ORF
413	6	AAU11070	Arabisdops	486	6	2.2	121	5	ABN90187	ABN90187 Human pol
414	6	ABR83556	Escherich	487	6	2.2	122	4	AAU42386	AAU42386 Propionib
415	6	ADC32781	Human nov	488	6	2.2	122	6	ABM38905	ABM38905 Propionib
416	6	ABG69875	Human sec	489	6	2.2	122	7	ADC88432	ADC88432 Ribosomal
417	6	ABP75798	Human sec	490	6	2.2	123	3	AAU48509	AAU48509 Arabisdops
418	6	AAU39605	Arabisdops	491	6	2.2	123	3	AAU41908	AAU41908 Arabisdops
419	6	ABP11069	Arabisdops	492	6	2.2	123	4	AAU17154	AAU17154 Peptide #
420	6	ABP77850	N. gonorr	493	6	2.2	123	4	ABK36155	ABK36155 Peptide #
421	6	AAU86838	Human imm	494	6	2.2	123	4	AAU29646	AAU29646 Peptide #
422	6	AAU66987	Propionib	495	6	2.2	123	4	ABK30964	ABK30964 Peptide #
423	6	ABG06348	Novel hum	496	6	2.2	123	4	ABK21542	ABK21542 Protein e
424	6	ABM63506	Propionib	497	6	2.2	123	4	AAU69324	AAU69324 Human bon
425	6	AAU40436	Propionib	498	6	2.2	123	4	AAU70858	AAU70858 Human bon
426	6	ABP31616	Human ORF	499	6	2.2	123	4	AAU56940	AAU56940 Human bra
427	6	ABM36955	Propionib	500	6	2.2	123	4	ABG51000	ABG51000 Human liv
428	6	AAU66991	Propionib	501	6	2.2	123	4	AAU04849	AAU04849 Peptide #
429	6	ABM63510	Propionib	502	6	2.2	123	5	ABG38941	ABG38941 Human pep
430	6	ABU35599	Protein e	503	6	2.2	125	3	AAU27567	AAU27567 Arabisdops
431	6	ADC88456	Ribosomal	504	6	2.2	125	3	AAU13561	AAU13561 Arabisdops
432	6	AAU74194	Human pro	505	6	2.2	125	3	AAU05365	AAU05365 Arabisdops
433	6	ABG25250	Arabisdops	506	6	2.2	125	4	AAU67588	AAU67588 Propionib
434	6	ABG26432	Novel hum	507	6	2.2	125	6	ABM64107	ABM64107 Propionib
435	6	ABBO6357	Human zin	508	6	2.2	126	3	AAU52850	AAU52850 Arabisdops
436	6	AAU4569	ORF1 prod	509	6	2.2	126	3	AAU52840	AAU52840 Arabisdops
437	6	AAU73871	Arabisdops	510	6	2.2	127	3	AAU48508	AAU48508 Arabisdops
438	6	AAU11068	Arabisdops	511	6	2.2	128	6	ABU23794	ABU23794 Protein e
439	6	AAU39604	Arabisdops	512	6	2.2	128	6	ABU24054	ABU24054 Protein e
440	6	ABP27936	Streptoco	513	6	2.2	130	3	AAU25249	AAU25249 Arabisdops
441	6	ABG93246	C. albica	514	6	2.2	131	2	AAU59719	AAU59719 Secreted
442	6	ABG93328	C. albica	515	6	2.2	131	4	ABG02740	ABG02740 Novel hum
443	6	ABP73948	Candida a	516	6	2.2	132	4	ABG27850	ABG27850 Novel hum
444	6	AAU50126	Propionib	517	6	2.2	132	5	ABP32084	ABP32084 Human ORF
445	6	ABBI5226	Human ner	518	6	2.2	132	6	ABU52320	ABU52320 Human GPC
446	6	ABBA49860	Listeria	519	6	2.2	133	2	AAU07877	AAU07877 Human sec
447	6	ABMA6645	Propionib	520	6	2.2	133	2	ABG03888	ABG03888 Novel hum
448	6	ADAA34275	Actinobac	521	6	2.2	133	2	AAU38829	AAU38829 Neisseria
449	6	AAU39333	Arabisdops	522	6	2.2	135	2	AAU38930	AAU38930 Neisseria
450	6	AAU61525	Sequence	523	6	2.2	135	2	AAU38927	AAU38927 Neisseria
451	6	AAU22712	zee may	524	6	2.2	135	3	AAU75021	AAU75021 Neisseria
452	6	ABBA4832	Drosophil	525	6	2.2	135	3	AAU75017	AAU75017 Neisseria
453	6	AAU63679	Human sec	526	6	2.2	135	3	AAU75020	AAU75020 Neisseria
454	6	AAU57672	Arabisdops	527	6	2.2	135	3	AAU57671	AAU57671 Arabisdops
455	6	AAU47868	Propionib	528	6	2.2	135	4	AAU03521	AAU03521 Human gen
456	6	AAU89900	C glutami	529	6	2.2	135	4	AAU23109	AAU23109 Novel hum
457	6	ABP31112	Human tra	530	6	2.2	135	4	AAU01323	AAU01323 Human pol
458	6	ABMA4387	Propionib	531	6	2.2	135	4	AAU44634	AAU44634 Propionib
459	6	ABM74341	DNA clone	532	6	2.2	135	5	ABG63423	ABG63423 Human alb
460	6	ABBA42766	Human ORF	533	6	2.2	135	6	ABP80909	ABP80909 N. gonorr
461	6	AAU29400	Human G P	534	6	2.2	135	6	ABM41153	ABM41153 Propionib
462	6	ABG60688	Novel G P	535	6	2.2	135	6	ABM77169	ABM77169 Stephyloc
463	6	ABB70709	Drosophil	536	6	2.2	136	3	AAU78979	AAU78979 Silkworm

537	6	2.2	137	4	AA012794	AA013794	Human pol	610	6	2.2	154	5	ABG79185	Abg79185	Human rib
538	6	2.2	137	5	ABP08419	ABP08419	Human CRF	611	6	2.2	155	2	AAy29570	AAy29570	Human lun
539	6	2.2	137	5	AAE25600	AAE25600	Human ino	612	6	2.2	155	3	AA844116	AA844116	Human can
540	6	2.2	138	4	AAO01261	AAO01261	Human pol	613	6	2.2	155	3	AA844451	AA844451	Human lun
541	6	2.2	138	4	AAU53569	AAU53569	Propionib	614	6	2.2	155	4	AAU65508	AAU65508	Human lun
542	6	2.2	138	6	ABM50088	ABM50088	Propionib	615	6	2.2	155	4	AAE13792	AAE13792	Human lun
543	6	2.2	139	3	AAG10263	AAG10263	Arabiidops	616	6	2.2	155	4	AA822274	AA822274	S. epider
544	6	2.2	139	3	AAG05364	AAG05364	Arabiidops	617	6	2.2	155	6	ABM62027	ABM62027	Propionib
545	6	2.2	139	3	AAG13560	AAG13560	Arabiidops	618	6	2.2	155	7	ADP65422	ADP65422	Human lun
546	6	2.2	139	3	AAG52935	AAG52935	Arabiidops	619	6	2.2	155	7	AD87676	AD87676	Human lun
547	6	2.2	139	3	AAG27566	AAG27566	Arabiidops	620	6	2.2	156	2	AA866743	AA866743	Biotin ca
548	6	2.2	139	4	AA876796	AA876796	Cornebacc	621	6	2.2	156	4	AAU36469	AAU36469	Pseudomon
549	6	2.2	139	4	AA880210	AA880210	Cornebacc	622	6	2.2	156	4	AAU42742	AAU42742	Human rep
550	6	2.2	139	5	ABG80342	ABG80342	C. glutam	623	6	2.2	156	4	ABG22975	ABG22975	Novel hum
551	6	2.2	140	3	AAG52949	AAG52949	Arabiidops	624	6	2.2	156	6	ABU38815	ABU38815	Protein e
552	6	2.2	140	3	AAG52939	AAG52939	Arabiidops	625	6	2.2	157	2	AAU38867	AAU38867	Neisseria
553	6	2.2	140	3	AAG52944	AAG52944	Arabiidops	626	6	2.2	157	3	ABG68060	ABG68060	Human int
554	6	2.2	140	5	ABP07556	ABP07556	Human CRF	627	6	2.2	158	4	ABG93070	ABG93070	C glutam
555	6	2.2	141	3	AA818483	AA818483	Zea may	628	6	2.2	159	4	ABG16754	ABG16754	Novel hum
556	6	2.2	141	6	ABO14990	ABO14990	Human NOV	629	6	2.2	159	5	AAU80548	AAU80548	Human G P
557	6	2.2	141	6	ABM64889	ABM64889	Propionib	630	6	2.2	159	6	AB849136	AB849136	Photocorb
558	6	2.2	141	7	ADC95085	ADC95085	E. fasciu	631	6	2.2	159	6	ABM69515	ABM69515	Listeria
559	6	2.2	142	4	AAU66857	AAU66857	Propionib	632	6	2.2	160	4	AAU49017	AAU49017	Propionib
560	6	2.2	142	4	ABG08438	ABG08438	Novel hum	633	6	2.2	160	6	ABM43536	ABM43536	Propionib
561	6	2.2	142	6	ABM63376	ABM63376	Propionib	634	6	2.2	161	6	AA860305	AA860305	Sequence
562	6	2.2	143	3	AA825805	AA825805	Arabiidops	635	6	2.2	161	4	AB855024	AB855024	Human sec
563	6	2.2	143	3	AA805522	AA805522	Arabiidops	636	6	2.2	161	5	ABP59105	ABP59105	Protein k
564	6	2.2	144	3	AA841907	AA841907	Arabiidops	637	6	2.2	161	5	AAU98692	AAU98692	Arabiidops
565	6	2.2	144	6	ABU48187	ABU48187	Protein e	638	6	2.2	161	5	AA886155	AA886155	S. pneumo
566	6	2.2	144	6	ABU47025	ABU47025	Protein e	639	6	2.2	163	3	AA857769	AA857769	Arabiidops
567	6	2.2	145	4	AAO07941	AAO07941	Human pol	640	6	2.2	163	4	AA818482	AA818482	Zea may
568	6	2.2	145	4	ABU31193	ABU31193	Protein e	641	6	2.2	164	4	ABG15828	ABG15828	Novel hum
569	6	2.2	145	6	ADB06440	ADB06440	Alloiococ	642	6	2.2	164	5	AB883046	AB883046	Translate
570	6	2.2	146	2	AAV59749	AAV59749	Human nor	643	6	2.2	164	5	AB883046	AB883046	Translate
571	6	2.2	146	3	AA825804	AA825804	Arabiidops	644	6	2.2	165	3	AA844980	AA844980	Human mut
572	6	2.2	146	4	AAU49172	AAU49172	Propionib	645	6	2.2	165	3	AA844980	AA844980	Human mut
573	6	2.2	146	4	AAU03698	AAU03698	Group B S	646	6	2.2	165	3	AAU44977	AAU44977	Human hyb
574	6	2.2	146	6	ABM45691	ABM45691	Propionib	647	6	2.2	165	4	AA855041	AA855041	Arabiidops
575	6	2.2	147	2	AAW98558	AAW98558	H. pylori	648	6	2.2	165	4	AA875335	AA875335	Human COI
576	6	2.2	147	5	AB870702	AB870702	Drosophi1	649	6	2.2	165	6	ABM69824	ABM69824	Photocorb
577	6	2.2	147	5	ABP38179	ABP38179	Staphyloc	650	6	2.2	165	6	ABU16597	ABU16597	Protein e
578	6	2.2	147	6	ABU26729	ABU26729	Protein e	651	6	2.2	166	1	AA830684	AA830684	Consensus
579	6	2.2	147	6	ADB06506	ADB06506	Alloiococ	652	6	2.2	166	1	AA830684	AA830684	Consensus
580	6	2.2	148	4	ABG29527	ABG29527	Novel hum	653	6	2.2	166	1	AA830683	AA830683	Synthetic
581	6	2.2	148	5	AB848719	AB848719	Listeria	654	6	2.2	166	1	AA840210	AA840210	Novel int
582	6	2.2	148	6	ABU01559	ABU01559	S. pneumo	655	6	2.2	166	1	AA850424	AA850424	Novel int
583	6	2.2	148	6	ABU32695	ABU32695	Protein e	656	6	2.2	166	1	AA860102	AA860102	Sequence
584	6	2.2	148	6	ABU40014	ABU40014	Protein e	657	6	2.2	166	1	AA860101	AA860101	Sequence
585	6	2.2	148	6	ABU41381	ABU41381	Protein e	658	6	2.2	166	1	AA860099	AA860099	Sequence
586	6	2.2	149	4	AAU65789	AAU65789	Propionib	659	6	2.2	166	1	AA860103	AA860103	Sequence
587	6	2.2	149	6	ABM62308	ABM62308	Propionib	660	6	2.2	166	1	AA860104	AA860104	Sequence
588	6	2.2	149	6	ABM65605	ABM65605	Propionib	661	6	2.2	166	1	AA860100	AA860100	Sequence
589	6	2.2	149	6	ABU14997	ABU14997	Protein e	662	6	2.2	166	1	AA890185	AA890185	Hybrid al
590	6	2.2	150	4	AAW95822	AAW95822	Human rep	663	6	2.2	166	1	AA890189	AA890189	Hybrid al
591	6	2.2	150	4	AB896353	AB896353	Human tes	664	6	2.2	166	1	AA890186	AA890186	Hybrid al
592	6	2.2	151	2	AAW62721	AAW62721	Streptococ	665	6	2.2	166	1	AA890186	AA890186	Hybrid al
593	6	2.2	151	4	AB850446	AB850446	Novel hum	666	6	2.2	166	1	AA890190	AA890190	Hybrid al
594	6	2.2	151	5	AB853410	AB853410	Lactococ	667	6	2.2	166	1	AA890187	AA890187	Hybrid al
595	6	2.2	151	6	ABM15935	ABM15935	Mycobacte	668	6	2.2	166	2	AA820053	AA820053	Human IFN
596	6	2.2	151	6	ABU52319	ABU52319	Human GPC	669	6	2.2	166	2	AA842814	AA842814	Lymphobia
597	6	2.2	151	6	ABU34492	ABU34492	Protein e	670	6	2.2	166	2	AA842812	AA842812	Lymphobia
598	6	2.2	151	6	ABU36955	ABU36955	Protein e	671	6	2.2	166	2	AA867762	AA867762	Interfero
599	6	2.2	151	7	AD884421	AD884421	Ribosomal	672	6	2.2	166	2	AA870088	AA870088	Recombina
600	6	2.2	152	6	ABU01766	ABU01766	S. pneumo	673	6	2.2	166	2	AA805337	AA805337	Alpha-int
601	6	2.2	152	6	ABU1960	ABU1960	Staphyloc	674	6	2.2	166	2	AA860141	AA860141	W. vaccae
602	6	2.2	153	4	AA878342	AA878342	Neutropept	675	6	2.2	166	2	AAU14888	AAU14888	Amino aci
603	6	2.2	153	4	AAU62113	AAU62113	Propionib	676	6	2.2	166	3	AAU44975	AAU44975	Human hyb
604	6	2.2	153	4	AAU03673	AAU03673	Group B S	677	6	2.2	166	3	AAU44983	AAU44983	Human mut
605	6	2.2	153	6	ABM58632	ABM58632	Propionib	678	6	2.2	166	3	AAU44974	AAU44974	Human hyb
606	6	2.2	153	6	ABU35768	ABU35768	Protein e	679	6	2.2	166	3	AAU44978	AAU44978	Human hyb
607	6	2.2	153	7	AD89419	AD89419	Ribosomal	680	6	2.2	166	3	AAU44978	AAU44978	Human hyb
608	6	2.2	154	7	AB85023	AB85023	Human sec	681	6	2.2	166	3	AAU44982	AAU44982	Human mut
609	6	2.2	154	4	AA878503	AA878503	Human SP-	682	6	2.2	166	3	AA828176	AA828176	Human int

683	6	2.2	166	3	AAB28177	Chimeric	756	6	2.2	167	4	AB559605	AB559605 Drosophila
684	6	2.2	166	3	AAV44834	Mutant by	757	6	2.2	167	4	ABG00676	ABG00676 Novel hum
685	6	2.2	166	3	AAV44832	Mutant by	758	6	2.2	167	6	AB537339	AB537339 Protein s
686	6	2.2	166	3	AAV44830	Hybrid in	759	6	2.2	167	6	ABU34732	ABU34732 Protein e
687	6	2.2	166	3	AAV44833	Mutant by	760	6	2.2	168	1	AA660318	AA660318 Sequence
688	6	2.2	166	3	AAV44827	Hybrid in	761	6	2.2	168	2	AA585866	AA585866 S. Pneumo
689	6	2.2	166	3	AAV44831	Hybrid in	762	6	2.2	168	3	AA557768	AA557768 Arbidops
690	6	2.2	166	3	AAV44829	Hybrid in	763	6	2.2	168	4	ABU41736	ABU41736 Propionib
691	6	2.2	166	3	AAV44835	Mutant by	764	6	2.2	168	4	ABG00677	ABG00677 Novel hum
692	6	2.2	166	3	AAV44826	Hybrid in	765	6	2.2	168	6	ABU00544	ABU00544 S. pneumo
693	6	2.2	166	3	AAV44826	Hybrid in	766	6	2.2	168	6	ABM38255	ABM38255 Propionib
694	6	2.2	166	3	AAV44826	Hybrid in	767	6	2.2	168	6	ABU29724	ABU29724 Protein e
695	6	2.2	166	4	AAV49154	Human int	768	6	2.2	170	1	AA660319	AA660319 Sequence
696	6	2.2	166	4	AAV49154	Human int	769	6	2.2	170	4	AA682328	AA682328 S. epider
697	6	2.2	166	4	AAV49159	Human int	770	6	2.2	170	4	AA690933	AA690933 Human col
698	6	2.2	166	4	AAV49163	Human int	771	6	2.2	170	5	ABU04304	ABU04304 Human col
699	6	2.2	166	4	AAV49162	Human int	772	6	2.2	170	5	ABP57955	ABP57955 Transposi
700	6	2.2	166	4	AAV49166	Human int	773	6	2.2	170	6	ABU01903	ABU01903 S. pneumo
701	6	2.2	166	4	AAV49166	Human int	774	6	2.2	170	7	ADC95570	ADC95570 E. faeciu
702	6	2.2	166	4	AAV49166	Human int	775	6	2.2	171	1	AA660764	AA660764 Plasmid e
703	6	2.2	166	4	AAV49166	Human int	776	6	2.2	171	6	AB61643	AB61643 Human int
704	6	2.2	166	4	AAV49166	Human int	777	6	2.2	171	6	AB622732	AB622732 Human alp
705	6	2.2	166	4	AAV49166	Human int	778	6	2.2	172	6	ADA35019	ADA35019 Actinocob
706	6	2.2	166	4	AAV49166	Human int	779	6	2.2	175	4	ABG17565	ABG17565 Novel hum
707	6	2.2	166	4	AAV49166	Human int	780	6	2.2	175	5	AAO15198	AAO15198 Rhodococc
708	6	2.2	166	4	AAV49166	Human int	781	6	2.2	175	5	AAU97143	AAU97143 Rhodococc
709	6	2.2	166	4	AAV49166	Human int	782	6	2.2	176	1	AA660557	AA660557 Interfero
710	6	2.2	166	4	AAV49166	Human int	783	6	2.2	176	6	AB118773	AB118773 Pseudomon
711	6	2.2	166	4	AAV49166	Human int	784	6	2.2	178	3	AA554807	AA554807 Arbidops
712	6	2.2	166	4	AAV49166	Human int	785	6	2.2	180	2	AA58436	AA58436 Streptoc
713	6	2.2	166	4	AAV49166	Human int	786	6	2.2	180	5	ABP69084	ABP69084 Human pol
714	6	2.2	166	4	AAV49166	Human int	787	6	2.2	182	2	AAW72889	AAW72889 Mycobacte
715	6	2.2	166	4	AAV49166	Human int	788	6	2.2	182	2	AAV21906	AAV21906 Amino aci
716	6	2.2	166	4	AAV49166	Human int	789	6	2.2	182	4	AAU17409	AAU17409 Novel sig
717	6	2.2	166	4	AAV49166	Human int	790	6	2.2	182	4	AAU18711	AAU18711 Renal and
718	6	2.2	166	4	AAV49166	Human int	791	6	2.2	182	6	ABP97326	ABP97326 Human pol
719	6	2.2	166	4	AAV49166	Human int	792	6	2.2	182	7	ABP94117	ABP94117 Human nov
720	6	2.2	166	4	AAV49166	Human int	793	6	2.2	182	7	ADC97014	ADC97014 E. faeciu
721	6	2.2	166	4	AAV49166	Human int	794	6	2.2	185	2	AAV34688	AAV34688 Chlamydia
722	6	2.2	166	4	AAV49166	Human int	795	6	2.2	185	3	ABG17774	ABG17774 Arbidops
723	6	2.2	166	4	AAV49166	Human int	796	6	2.2	185	4	ABG01511	ABG01511 Novel hum
724	6	2.2	166	4	AAV49166	Human int	797	6	2.2	186	4	AAV43545	AAV43545 Human pol
725	6	2.2	166	4	AAV49166	Human int	798	6	2.2	186	4	AAV92477	AAV92477 C glutami
726	6	2.2	166	4	AAV49166	Human int	799	6	2.2	187	2	AAV62368	AAV62368 Interfero
727	6	2.2	166	4	AAV49166	Human int	800	6	2.2	187	2	AAV54038	AAV54038 Human pan
728	6	2.2	166	4	AAV49166	Human int	801	6	2.2	188	1	AAV20009	AAV20009 Hybrid hu
729	6	2.2	166	4	AAV49166	Human int	802	6	2.2	188	3	AAV42560	AAV42560 Human ORF
730	6	2.2	166	4	AAV49166	Human int	803	6	2.2	189	1	AAV20108	AAV20108 Sequence
731	6	2.2	166	4	AAV49166	Human int	804	6	2.2	189	1	AAV20104	AAV20104 Sequence
732	6	2.2	166	4	AAV49166	Human int	805	6	2.2	189	1	AAV20105	AAV20105 Sequence
733	6	2.2	166	4	AAV49166	Human int	806	6	2.2	189	1	AAV20111	AAV20111 Sequence
734	6	2.2	166	4	AAV49166	Human int	807	6	2.2	189	1	AAV20112	AAV20112 Sequence
735	6	2.2	166	4	AAV49166	Human int	808	6	2.2	189	1	AAV30228	AAV30228 Sequence
736	6	2.2	166	4	AAV49166	Human int	809	6	2.2	189	1	AAV30179	AAV30179 Sequence
737	6	2.2	166	4	AAV49166	Human int	810	6	2.2	189	1	AAV30183	AAV30183 Sequence
738	6	2.2	166	4	AAV49166	Human int	811	6	2.2	189	1	AAV30164	AAV30164 Sequence
739	6	2.2	166	4	AAV49166	Human int	812	6	2.2	189	1	AAV30165	AAV30165 Sequence
740	6	2.2	166	4	AAV49166	Human int	813	6	2.2	189	1	AAV30101	AAV30101 Sequence
741	6	2.2	166	4	AAV49166	Human int	814	6	2.2	189	1	AAV40741	AAV40741 Sequence
742	6	2.2	166	4	AAV49166	Human int	815	6	2.2	189	1	AAV40179	AAV40179 Sequence
743	6	2.2	166	4	AAV49166	Human int	816	6	2.2	189	1	AAV40123	AAV40123 Sequence
744	6	2.2	166	4	AAV49166	Human int	817	6	2.2	189	1	AAV40127	AAV40127 Sequence
745	6	2.2	166	4	AAV49166	Human int	818	6	2.2	189	1	AAV50306	AAV50306 Human int
746	6	2.2	166	4	AAV49166	Human int	819	6	2.2	189	1	AAV50104	AAV50104 Sequence
747	6	2.2	166	4	AAV49166	Human int	820	6	2.2	189	2	AAV07680	AAV07680 IFN-alpha
748	6	2.2	166	4	AAV49166	Human int	821	6	2.2	189	2	AAV64240	AAV64240 Human int
749	6	2.2	166	4	AAV49166	Human int	822	6	2.2	189	2	AAV62452	AAV62452 Human int
750	6	2.2	166	4	AAV49166	Human int	823	6	2.2	189	2	AAV70373	AAV70373 Human int
751	6	2.2	166	4	AAV49166	Human int	824	6	2.2	189	2	AAV94083	AAV94083 Human int
752	6	2.2	166	4	AAV49166	Human int	825	6	2.2	189	3	AAV59395	AAV59395 Human int
753	6	2.2	166	4	AAV49166	Human int	826	6	2.2	189	3	AAV12867	AAV12867 Human int
754	6	2.2	166	4	AAV49166	Human int	827	6	2.2	189	3	AAV15533	AAV15533 Consensus
755	6	2.2	166	4	AAV49166	Human int	828	6	2.2	189	4	AAV49779	AAV49779 Human int

829	6	2.2	189	4	AA849780	Aab49780	Human	int	902	6	2.2	206	3	AA645097	Aag45097	Arabidops
830	6	2.2	189	4	AAE14364	Aae14364	Human	int	903	6	2.2	207	4	AA607069	Aab60769	Gene 18 r
831	6	2.2	189	5	AAU84283	Aau84283	Human	end	904	6	2.2	207	4	AA607770	AA607770	Gene 18 r
832	6	2.2	189	5	ABG68066	ABg68066	Human	int	905	6	2.2	207	6	ABU17138	ABu17138	Protein e
833	6	2.2	189	5	ABG68061	ABg68061	Human	int	906	6	2.2	208	4	AA692507	AAg92507	C glutami
834	6	2.2	189	5	ABG68075	ABg68075	Human	int	907	6	2.2	209	4	AAU34087	AAu34087	Staphyloc
835	6	2.2	189	5	ABG68076	ABg68076	Human	int	908	6	2.2	209	4	AAU36374	AAu36374	Pseudomon
836	6	2.2	189	5	ABG68063	ABg68063	Human	int	909	6	2.2	209	4	AAU36374	AAu36374	Staphyloc
837	6	2.2	189	5	ABG68059	ABg68059	Human	int	909	6	2.2	209	4	AAU36374	AAu36374	Pseudomon
838	6	2.2	189	5	ABG68068	ABg68068	Human	int	910	6	2.2	209	6	AAU36374	AAu36374	Staphyloc
839	6	2.2	189	5	ABG68067	ABg68067	Human	int	911	6	2.2	209	6	AAU36374	AAu36374	Staphyloc
840	6	2.2	189	5	ABG68073	ABg68073	Human	int	912	6	2.2	210	6	ABU27599	ABu27599	Protein e
841	6	2.2	189	5	ABG68072	ABg68072	Human	int	913	6	2.2	211	3	AAU17773	AAu17773	Arabidops
842	6	2.2	189	5	ABG68071	ABg68071	Human	int	914	6	2.2	211	4	AAU00776	AAu00776	Human
843	6	2.2	189	5	ABG68062	ABg68062	Human	int	915	6	2.2	212	6	ABU31486	ABu31486	Protein e
844	6	2.2	189	5	ABG68065	ABg68065	Human	int	916	6	2.2	213	6	AAU34129	AAu34129	Acinetoba
845	6	2.2	189	5	ABG68074	ABg68074	Human	int	917	6	2.2	213	6	AAU23545	AAu23545	Mycoplasma
846	6	2.2	189	5	ABG68076	ABg68076	Human	int	918	6	2.2	213	3	AAU43037	AAu43037	Human
847	6	2.2	189	5	ABG68076	ABg68076	Human	int	919	6	2.2	213	4	ABG00998	ABg00998	Novel hum
848	6	2.2	189	5	ABG68076	ABg68076	Human	int	920	6	2.2	213	6	ABU16388	ABu16388	Protein e
849	6	2.2	189	5	ABG68076	ABg68076	Human	int	921	6	2.2	214	5	ABP26854	ABp26854	Streptoco
850	6	2.2	189	5	ABG68076	ABg68076	Human	int	922	6	2.2	214	5	ABP73526	ABp73526	Candida
851	6	2.2	189	5	ABG68076	ABg68076	Human	int	923	6	2.2	214	6	ABG69016	ABg69016	Photocro
852	6	2.2	189	5	ABG68076	ABg68076	Human	int	924	6	2.2	215	5	ABP26910	ABp26910	Streptoco
853	6	2.2	189	5	ABG68076	ABg68076	Human	int	925	6	2.2	215	5	AAE23611	AAe23611	Bacillus
854	6	2.2	189	5	ABG68076	ABg68076	Human	int	926	6	2.2	215	6	ABR53146	ABr53146	Protein s
855	6	2.2	189	5	ABG68076	ABg68076	Human	int	927	6	2.2	216	5	ABG93145	ABg93145	S. cerevi
856	6	2.2	189	5	ABG68076	ABg68076	Human	int	928	6	2.2	216	6	ABU23576	ABu23576	Protein e
857	6	2.2	189	5	ABG68076	ABg68076	Human	int	929	6	2.2	216	6	ABU26095	ABu26095	Aspergill
858	6	2.2	189	5	ABG68076	ABg68076	Human	int	930	6	2.2	217	2	AAU23469	AAu23469	Recombina
859	6	2.2	189	5	ABG68076	ABg68076	Human	int	931	6	2.2	217	2	AAU23469	AAu23469	Recombina
860	6	2.2	189	5	ABG68076	ABg68076	Human	int	932	6	2.2	217	2	AAU23469	AAu23469	Recombina
861	6	2.2	189	5	ABG68076	ABg68076	Human	int	933	6	2.2	217	2	AAU23469	AAu23469	Recombina
862	6	2.2	189	5	ABG68076	ABg68076	Human	int	934	6	2.2	217	2	AAU23469	AAu23469	Recombina
863	6	2.2	189	5	ABG68076	ABg68076	Human	int	935	6	2.2	217	2	AAU23469	AAu23469	Recombina
864	6	2.2	189	5	ABG68076	ABg68076	Human	int	936	6	2.2	217	2	AAU23469	AAu23469	Recombina
865	6	2.2	189	5	ABG68076	ABg68076	Human	int	937	6	2.2	217	2	AAU23469	AAu23469	Recombina
866	6	2.2	189	5	ABG68076	ABg68076	Human	int	938	6	2.2	217	2	AAU23469	AAu23469	Recombina
867	6	2.2	189	5	ABG68076	ABg68076	Human	int	939	6	2.2	217	2	AAU23469	AAu23469	Recombina
868	6	2.2	189	5	ABG68076	ABg68076	Human	int	940	6	2.2	217	2	AAU23469	AAu23469	Recombina
869	6	2.2	189	5	ABG68076	ABg68076	Human	int	941	6	2.2	217	6	ABP81353	ABp81353	Streptoco
870	6	2.2	189	5	ABG68076	ABg68076	Human	int	942	6	2.2	217	7	ADP66458	ADp66458	Human
871	6	2.2	189	5	ABG68076	ABg68076	Human	int	943	6	2.2	217	7	ADP66458	ADp66458	Human
872	6	2.2	189	5	ABG68076	ABg68076	Human	int	944	6	2.2	217	7	ADP66458	ADp66458	Human
873	6	2.2	189	5	ABG68076	ABg68076	Human	int	945	6	2.2	218	3	AAU39423	AAu39423	Arabidops
874	6	2.2	189	5	ABG68076	ABg68076	Human	int	946	6	2.2	218	6	AAU39423	AAu39423	Arabidops
875	6	2.2	189	5	ABG68076	ABg68076	Human	int	947	6	2.2	219	2	AAU39423	AAu39423	Arabidops
876	6	2.2	189	5	ABG68076	ABg68076	Human	int	948	6	2.2	219	2	AAU39423	AAu39423	Arabidops
877	6	2.2	189	5	ABG68076	ABg68076	Human	int	949	6	2.2	219	2	AAU39423	AAu39423	Arabidops
878	6	2.2	189	5	ABG68076	ABg68076	Human	int	950	6	2.2	219	4	AAU39423	AAu39423	Arabidops
879	6	2.2	189	5	ABG68076	ABg68076	Human	int	951	6	2.2	220	4	AAU39423	AAu39423	Arabidops
880	6	2.2	189	5	ABG68076	ABg68076	Human	int	952	6	2.2	220	6	AAU39423	AAu39423	Arabidops
881	6	2.2	189	5	ABG68076	ABg68076	Human	int	953	6	2.2	221	2	AAU39423	AAu39423	Arabidops
882	6	2.2	189	5	ABG68076	ABg68076	Human	int	954	6	2.2	221	2	AAU39423	AAu39423	Arabidops
883	6	2.2	189	5	ABG68076	ABg68076	Human	int	955	6	2.2	221	3	AAU39423	AAu39423	Arabidops
884	6	2.2	189	5	ABG68076	ABg68076	Human	int	956	6	2.2	221	3	AAU39423	AAu39423	Arabidops
885	6	2.2	189	5	ABG68076	ABg68076	Human	int	957	6	2.2	221	5	AAU39423	AAu39423	Arabidops
886	6	2.2	189	5	ABG68076	ABg68076	Human	int	958	6	2.2	221	7	AAU39423	AAu39423	Arabidops
887	6	2.2	189	5	ABG68076	ABg68076	Human	int	959	6	2.2	222	4	AAU39423	AAu39423	Arabidops
888	6	2.2	189	5	ABG68076	ABg68076	Human	int	960	6	2.2	222	6	AAU39423	AAu39423	Arabidops
889	6	2.2	189	5	ABG68076	ABg68076	Human	int	961	6	2.2	222	6	AAU39423	AAu39423	Arabidops
890	6	2.2	189	5	ABG68076	ABg68076	Human	int	962	6	2.2	223	3	AAU39423	AAu39423	Arabidops
891	6	2.2	189	5	ABG68076	ABg68076	Human	int	963	6	2.2	223	3	AAU39423	AAu39423	Arabidops
892	6	2.2	189	5	ABG68076	ABg68076	Human	int	964	6	2.2	223	5	AAU39423	AAu39423	Arabidops
893	6	2.2	189	5	ABG68076	ABg68076	Human	int	965	6	2.2	224	5	AAU39423	AAu39423	Arabidops
894	6	2.2	189	5	ABG68076	ABg68076	Human	int	966	6	2.2	224	5	AAU39423	AAu39423	Arabidops
895	6	2.2	189	5	ABG68076	ABg68076	Human	int	967	6	2.2	225	4	AAU39423	AAu39423	Arabidops
896	6	2.2	189	5	ABG68076	ABg68076	Human	int	968	6	2.2	226	4	AAU39423	AAu39423	Arabidops
897	6	2.2	189	5	ABG68076	ABg68076	Human	int	969	6	2.2	226	4	AAU39423	AAu39423	Arabidops
898	6	2.2	189	5	ABG68076	ABg68076	Human	int	970	6	2.2	227	6	AAU39423	AAu39423	Arabidops
899	6	2.2	189	5	ABG68076	ABg68076	Human	int	971	6	2.2	228	3	AAU39423	AAu39423	Arabidops
900	6	2.2	189	5	ABG68076	ABg68076	Human	int	972	6	2.2	228	3	AAU39423	AAu39423	Arabidops
901	6	2.2	189	5	ABG68076	ABg68076	Human	int	973	6	2.2	228	3	AAU39423	AAu39423	Arabidops
									974			228	3	AAU39423	AAu39423	Arabidops

975 6 2.2 228 6 ABU22153
976 6 2.2 229 3 AAG52948
977 6 2.2 229 6 ABM69350
978 6 2.2 229 6 ADK6158
979 6 2.2 230 4 AAU23381
980 6 2.2 230 4 AAG91949
981 6 2.2 231 4 AAG75438
982 6 2.2 231 4 ABG26456
983 6 2.2 232 4 AAM34978
984 6 2.2 232 4 ABM61621
985 6 2.2 233 2 AAR12363
986 6 2.2 234 2 AAR42265
987 6 2.2 234 3 AAG49817
988 6 2.2 234 5 ABP28742
989 6 2.2 234 6 ABP46706
990 6 2.2 235 4 ABU39508
991 6 2.2 236 4 AAG32550
992 6 2.2 236 4 ABG15597
993 6 2.2 237 2 AAY59839
994 6 2.2 237 2 AAY74002
995 6 2.2 237 3 AAG46706
996 6 2.2 237 3 AAG19425
997 6 2.2 237 3 AAG19438
998 6 2.2 237 3 AAG46702
999 6 2.2 237 3 AAG19579
1000 6 2.2 237 3 AAG46689

ALIGNMENTS

RESULT 1
AAB20104 standard; protein; 276 AA.

XX AAB20104;
XX 23-APR-2001 (first entry)
XX Moraxella catarrhalis BASB11 protein.
XX DE
XX BASB11; infection; otitis media; pneumonia; diagnosis; therapy;
XX KW antibacterial; antimicrobial; vaccine.
XX OS
XX Moraxella catarrhalis.
XX PN MO200100837-A1.
XX PD 04-JAN-2001.
XX PF 23-JUN-2000; 2000WO-EP005852.
XX PR 25-JUN-1999; 99GB-00014945.
XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX PI Thonard J;
XX WPI: 2001-123013/13.
XX DR N-PSDB; AAF30040.
XX PT New BASB11 polypeptides of Moraxella catarrhalis useful for diagnostic,
XX PT prophylactic and therapeutic purposes against microbial diseases,
XX PT preferably bacterial infections.
XX Claim 1; Page 63; 79pp; English.
XX The present sequence is that of BASB11 protein from Moraxella
XX catarrhalis strain MC2931 (ATCC 43617), a causative agent of otitis media
XX in children and pneumonia in adults. The invention provides BASB11
XX polypeptides, and polynucleotides encoding them, as well as expression
XX vectors, host cells and methods for producing BASB11 polypeptides using
XX recombinant methods. Also claimed is a vaccine composition comprising a

CC BASB11 polypeptide, an immunogenic fragment of a BASB11 polypeptide, or
CC a polypeptide having at least 85% amino acid sequence identity to
CC BASB11, or comprising a polynucleotide encoding such a polypeptide. A
CC claimed method of diagnosing a Moraxella infection involves identifying a
CC BASB11 polypeptide or antibody. A claimed therapeutic composition useful
CC in treating humans with M. catarrhalis infection comprises at least 1
CC antibody directed against a BASB11 polypeptide. BASB11 polypeptides
CC also have utility in raising specific antibodies, and in screening for
CC antibacterial drugs

Seq Sequence 276 AA;
Query Match 100.0%; Score 276; DB 4; Length 276;
Best Local Similarity 100.0%; Pred. No. 9,3e-251; Indels 0; Gaps 0;
Matches 276; Conservative 0; Mismatches 0;

QY 1 MNPFGKINGICATAGSIALAGCSNOSNEPAAISKTAAQTIKYGVAGCEQAVAEVAGVAK 60
DB 1 MNPFGKINGICATAGSIALAGCSNOSNEPAAISKTAAQTIKYGVAGCEQAVAEVAGVAK 60
QY 61 EKNVITVELVFNPDYAMNSAVSGEIDANNMOKPYLEKDSQSKGLNNTVYVNTFVYP 120
DB 61 EKNVITVELVFNPDYAMNSAVSGEIDANNMOKPYLEKDSQSKGLNNTVYVNTFVYP 120
QY 121 LAGYSTKIKTLNELKDGATIAVPNDPSNLARALILLEKQGLIKLKDNTNLFSTLIDIVEN 180
DB 121 LAGYSTKIKTLNELKDGATIAVPNDPSNLARALILLEKQGLIKLKDNTNLFSTLIDIVEN 180
QY 181 PKGLVKEVDTSVAAARAIDVDLA VVNNNTYAGVGLTASENGVEVEDKSPYNTIYARA 240
DB 181 PKGLVKEVDTSVAAARAIDVDLA VVNNNTYAGVGLTASENGVEVEDKSPYNTIYARA 240
QY 241 DNKOSKAIQDFVKA VQYDVEAEAKKQFKDGVITGM 276
DB 241 DNKOSKAIQDFVKA VQYDVEAEAKKQFKDGVITGM 276

RESULT 2
AAO17582
ID AAO17582 standard; protein; 276 AA.
XX AAO17582;
XX 19-JUL-2002 (first entry)
XX DE M catarrhalis MCA101896 protein SEQ ID NO: 44.
XX KW Moraxella; vaccine; respiratory tract infection; antiinflammatory;
XX KW auditory; antibacterial; otitis media; sinusitis; pneumonia.
XX OS
XX Moraxella catarrhalis.
XX PN MO200218595-A2.
XX PD 07-MAR-2002.
XX PF 28-AUG-2001; 2001WO-CA001221.
XX PR 28-AUG-2000; 2000US-0228294P.
XX PR 28-AUG-2000; 2000US-0228295P.
XX PR 28-AUG-2000; 2000US-0228296P.
XX PR 29-AUG-2000; 2000US-0228438P.
XX PR 29-AUG-2000; 2000US-0228439P.
XX PR 29-AUG-2000; 2000US-0228440P.
XX PR 29-AUG-2000; 2000US-0228441P.
XX PR 29-AUG-2000; 2000US-0228442P.
XX PR 29-AUG-2000; 2000US-0228443P.
XX PR 29-AUG-2000; 2000US-0228511P.
XX PR 29-AUG-2000; 2000US-0228512P.
XX PR 29-AUG-2000; 2000US-0228742P.
XX PR 29-AUG-2000; 2000US-0228773P.
XX PR 01-SEP-2000; 2000US-0229465P.
XX PR 01-SEP-2000; 2000US-0229474P.

PR 01-SEP-2000; 2000US-0229475P.
 PR 01-SEP-2000; 2000US-0229478P.
 PR 05-SEP-2000; 2000US-0229740P.
 PR 05-SEP-2000; 2000US-0229803P.
 PR 05-SEP-2000; 2000US-0229804P.
 PR 05-SEP-2000; 2000US-0229805P.
 PR 05-SEP-2000; 2000US-0229806P.
 PR 05-SEP-2000; 2000US-0229809P.
 PR 05-SEP-2000; 2000US-0229811P.
 PR 06-SEP-2000; 2000US-0230214P.
 PR 06-SEP-2000; 2000US-0230250P.
 PR 06-SEP-2000; 2000US-0230252P.
 PA (AVET) AVENTIS PASTEUR LTD.
 PI Loomore S, Wang J, Bradley B, Ochs M, Yang Y;
 PI WPI; 2002-404555/43.
 DR N-PSDB; AAL46514.
 DR WPI; 2002-404555/43.
 XX Moraxella polypeptide and polynucleotides useful as vaccine for
 PT immunizing a host e.g. humans against disease e.g. otitis media,
 PT pneumonia, caused by infection of the bacteria.
 XX Claim 28; Fig 43; 277pp; English.
 PS The present invention provides the protein and coding sequences of
 CC proteins from Moraxella catarrhalis. These can be used to produce
 CC vaccines which protect against M. catarrhalis infection, which can cause
 CC otitis media, respiratory infection, sinusitis, and pneumonia. The
 CC present sequence is a protein of the invention
 XX Sequence 276 AA;
 SQ
 Query Match 88.0%; Score 243; DB 5; Length 276;
 Best Local Similarity 100.0%; Pred. No. 9.4e-220; Indels 0; Gaps 0;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 34 TAAQTIRKGVAGPEQVAEVAQVAKKYNLTVELVEFNDYAMPNSAVSKGELDANAMQ 93
 DB 34 TAAQTIRKGVAGPEQVAEVAQVAKKYNLTVELVEFNDYAMPNSAVSKGELDANAMQ 93
 QY 94 HKPYLEKDSQEKGLNLTIVGNFVPLAGYSTKIKTLNELKDGATIAVNDPSNLARAL 153
 DB 94 HKPYLEKDSQEKGLNLTIVGNFVPLAGYSTKIKTLNELKDGATIAVNDPSNLARAL 153
 QY 154 ILEKQGLIKLKONTNLFTTLDIVENPKKLVKEVDTSVAARALDDVDLAVVNNNYAQO 213
 DB 154 ILEKQGLIKLKONTNLFTTLDIVENPKKLVKEVDTSVAARALDDVDLAVVNNNYAQO 213
 QY 214 VGLTASENVFVEDKSPYVNIIVARADKSKALIDFVKAYQTDVEVEAKKQKRDGYI 273
 DB 214 VGLTASENVFVEDKSPYVNIIVARADKSKALIDFVKAYQTDVEVEAKKQKRDGYI 273
 QY 274 KGW 276
 DB 274 KGW 276
 RESULT 3
 ABUS3462
 ID ABUS3462 standard; protein; 276 AA.
 AC ABUS3462;
 DT 19-JUN-2003 (first entry)
 DE Protein encoded by Prokaryotic essential gene #20989.
 XX Anti-sense; prokaryotic essential gene; cell proliferation; drug design.
 XX Moraxella catarrhalis.

PN MO200277183-A2.
 XX 03-OCT-2002.
 PD 21-MAR-2002; 2002WO-US009107.
 XX 21-MAR-2002; 2001US-00815242.
 XX 21-MAR-2002; 2001US-00948993.
 XX 06-SEP-2001; 2001US-0342823P.
 PR 25-OCT-2001; 2001US-00072851.
 PR 08-FEB-2002; 2002US-0362699P.
 PR 06-MAR-2002; 2002US-0362699P.
 XX (ELIT) ELITRA PHARM INC.
 PA Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 PI WPI; 2003-029926/02.
 DR N-PSDB; ACA39332.
 DR WPI; 2003-029926/02.
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX Claim 25; SEQ ID NO 63386; 176pp; English.
 PS The invention relates to an isolated nucleic acid comprising any one of
 CC the 623 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 276 AA;
 SQ
 Query Match 88.0%; Score 243; DB 6; Length 276;
 Best Local Similarity 100.0%; Pred. No. 9.4e-220; Indels 0; Gaps 0;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 34 TAAQTIRKGVAGPEQVAEVAQVAKKYNLTVELVEFNDYAMPNSAVSKGELDANAMQ 93
 DB 34 TAAQTIRKGVAGPEQVAEVAQVAKKYNLTVELVEFNDYAMPNSAVSKGELDANAMQ 93
 QY 94 HKPYLEKDSQEKGLNLTIVGNFVPLAGYSTKIKTLNELKDGATIAVNDPSNLARAL 153
 DB 94 HKPYLEKDSQEKGLNLTIVGNFVPLAGYSTKIKTLNELKDGATIAVNDPSNLARAL 153
 QY 154 ILEKQGLIKLKONTNLFTTLDIVENPKKLVKEVDTSVAARALDDVDLAVVNNNYAQO 213

CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 273 AA;

Query Match 6.9%; Score 19; DB 6; Length 273;
Best Local Similarity 100.0%; Pred. No. 2.7e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 LNNLVIGNTFVYPLAGYS 125
Db 104 LNNLVIGNTFVYPLAGYS 122

RESULT 9
AB076094 ABB76094 standard; peptide; 20 AA.

XX ABB76094;

DT 15-JUL-2002 (first entry)

XX Haemophilus influenzae BASB202 protein T-helper cell epitope.

XX BASB202; NTH1; infection; vaccine; genetic immunisation; auditory;
KM anti-inflammation; antibacterial; immunostimulant; otitis media;
XX T lymphocyte; epitope.

OS Haemophilus influenzae.

XX WO200224729-A2.

XX 28-MAR-2002.

XX 18-SEP-2001; 2001WO-EP010979.

XX 19-SEP-2000; 2000GB-00022392.

XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

XX Thonard J;

XX WPI; 2002-383180/41.

XX New isolated BASB202 polypeptide of nontypeable Haemophilus influenzae
PT useful for diagnosing a disease and in generating an immune response in
XX an animal.

XX Example 13; Page 72; 90pp; English.

XX The present sequence is a potential T-helper cell epitope of the BASB202
CC protein of nontypeable Haemophilus influenzae (NTH1). It corresponds to
CC amino acid residues 107-126 of the BASB202 sequences given in ABB76073
CC and ABB76074. T-helper cell epitopes are peptides bound to human
CC leukocyte antigen (HLA) class II molecules and recognised by T-helper
CC cells. Peptides comprising BASB202 T-cell epitopes (see ABB76090-99) can
CC be useful (preferably when conjugated to a peptide, polypeptide or
CC polyaacharide) in a vaccine composition for the prevention of NTH1-
CC related diseases, such as otitis media, pneumonia, sinusitis, nosocomial
CC infections, invasive disease, chronic otitis media with hearing loss,
CC fluid accumulation in the middle ear, auditory nerve damage, delayed
CC speech learning, infection of the upper respiratory tract, and
CC inflammation of the middle ear

XX Sequence 20 AA;

Query Match 5.8%; Score 16; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.9e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 110 LVIVGNTFVYPLAGYS 125
Db 1 LVIVGNTFVYPLAGYS 16

RESULT 10
ABU40467

XX ABU40467 standard; protein; 271 AA.

XX ABU40467;

DT 19-JUN-2003 (first entry)

XX Protein encoded by prokaryotic essential gene #25994.

XX Antisense, prokaryotic essential gene; cell proliferation; drug design.

XX Proteus sp.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

XX 25-OCT-2001; 2001US-0342923P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362659P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

XX N-PSDB; ACA44337.

XX Claim 25; SEQ ID NO 68391; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this

CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 271 AA;

Query Match 5.8%; Score 16; DB 6; Length 271;

Best Local Similarity 100.0%; Pred. No. 1.8e-06;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

224 FVBDKSPYNIIVAR 239

219 FVBDKSPYNIIVAR 234

RESULT 11

ABB76095 standard; peptide; 20 AA.

ABB76095;

15-JUL-2002 (first entry)

Haemophilus influenzae BASB202 protein T-helper cell epitope.

BASB202; NTH1; infection; vaccine; genetic immunisation; auditory;

antimicrobial; antibacterial; immunostimulant; otitis media;

T lymphocyte; epitope.

Haemophilus influenzae.

WO200224729-A2.

28-MAR-2002.

18-SEP-2001; 2001WO-EP010979.

19-SEP-2000; 2000GB-00022992.

(SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

Thomard J;

WPI; 2002-383180/41.

New isolated BASB202 polypeptide of nontypeable Haemophilus influenzae

useful for diagnosing a disease and in generating an immune response in

an animal.

Example 13; Page 72; 90pp; English.

The present sequence is a potential T-helper cell epitope of the BASB202

protein of nontypeable Haemophilus influenzae (NTH1). It corresponds to

amino acid residues 150-159 of the BASB202 sequences given in ABB76073

and ABB76074. T-helper cell epitopes are peptides bound to human

leukocyte antigen (HLA) class II molecules and recognised by T-helper

cells. Peptides comprising BASB202 T-cell epitopes (see ABB76090-99) can

be useful (preferably when conjugated to a peptide, polypeptide or

polysaccharide) in a vaccine composition for the prevention of NTH1-

related diseases, such as otitis media, pneumonia, sinusitis, nosocomial

infections, invasive disease chronic otitis media with hearing loss,

fluid accumulation in the middle ear, auditory nerve damage, delayed

speech learning, infection of the upper respiratory tract, and

inflammation of the middle ear

Sequence 20 AA;

Query Match 5.1%; Score 14; DB 5; Length 20;

Best Local Similarity 100.0%; Pred. No. 1.5e-05;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

153 LILKOGILKMD 166

|||||

DB 1 LILKOGILKMD 14

RESULT 12

ABU24914 standard; protein; 270 AA.

ABU24914;

19-JUN-2003 (first entry)

Protein encoded by prokaryotic essential gene #10441.

Antisense; prokaryotic essential gene; cell proliferation; drug design.

Clostridium botulinum.

WO20027183-A2.

03-OCT-2002.

21-MAR-2002; 2002WO-US009107.

21-MAR-2001; 2001US-00815242.

06-SEP-2001; 2001US-00948993.

25-OCT-2001; 2001US-0342923P.

08-FEB-2002; 2002US-00072851.

06-MAR-2002; 2002US-0362699P.

(ELT-) ELITRA PHARM INC.

Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

WPI; 2003-029926/02.

N-PsDB; ACA28784.

New antisense nucleic acid, useful for identifying proteins or screening

for homologous nucleic acids required for cellular proliferation to

isolate candidate molecules for rational drug discovery programs.

Claim 25; SEQ ID NO 52838; 1766pp; English.

The invention relates to an isolated nucleic acid comprising any one of

the 6213 antisense sequences given in the specification where expression

of the nucleic acid inhibits proliferation of a cell. Also included are:

(1) a vector comprising a promoter operably linked to the nucleic acid

encoding a polypeptide whose expression is inhibited by the antisense

nucleic acid; (2) a host cell containing the vector; (3) an isolated

polypeptide or its fragment whose expression is inhibited by the

antisense nucleic acid; (4) an antibody capable of specifically binding

the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

proliferation or the activity of a gene in an operon required for

proliferation; (7) identifying a compound that influences the activity of

the gene product or that has an activity against a biological pathway

required for proliferation, or that inhibits cellular proliferation; (8)

identifying a gene required for cellular proliferation or the biological

pathway in which a proliferation-required gene or its gene product lies

or a gene on which the test compound that inhibits proliferation of an

organism acts; (9) manufacturing an antibiotic; (10) profiling a

compound's activity; (11) a culture comprising strains in which the gene

product is overexpressed or underexpressed; (12) determining the extent

to which each of the strains is present in a culture or collection of

strains; or (13) identifying the target of a compound that inhibits the

proliferation of an organism. The antisense nucleic acids are useful for

identifying proteins or screening for homologous nucleic acids required

for cellular proliferation to isolate candidate molecules for rational

drug discovery programs, or for screening homologous nucleic acids

required for proliferation in cells other than S. aureus, S. typhimurium,

K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of

the target prokaryotic essential genes. Note: The sequence data for this

patent did not form part of the printed specification, but was obtained

in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 270 AA;
 Query Match 4.7%; Score 13; DB 6; Length 270;
 Best Local Similarity 100.0%; Pred. No. 0.0012;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 134 LKDGATIAVPNDP 146
 |||||
 Db 130 LKDGATIAVPNDP 142
 RESULT 13
 ABU39463
 ID ABU39463 standard; protein; 276 AA.
 XX
 AC ABU39463;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #24990.
 XX
 KM Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 OS Pasteurella multocida.
 XX
 PN WO200277183-A2.
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELITR) ELITRA PHARM INC.
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KU, Zyskind JW;
 PI Wall D, Trawick JD, Carr GF, Yamamoto R, Forsyth RA, Xu HH;
 XX N-PSDB; ACA43333.
 DR WPI: 2003-029936/02.
 DR
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 67387; 1766pp; English.
 CC
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of

CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 276 AA;
 Query Match 4.3%; Score 12; DB 6; Length 276;
 Best Local Similarity 100.0%; Pred. No. 0.01;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 155 LLEKQGLIKXD 166
 |||||
 Db 155 LLEKQGLIKXD 166
 RESULT 14
 AAB70412
 ID AAB70412 standard; protein; 67 AA.
 XX
 AC AAB70412;
 XX
 DT 02-MAY-2001 (first entry)
 XX
 DE A. actinomycetemcomitans immunoreactive protein SEQ ID NO:15.
 XX
 KM Actinobacillus actinomycetemcomitans; microbial; infection; vaccine;
 KM identification; localised juvenile periodontitis; antibacterial;
 KM antiinflammatory.
 XX
 OS Actinobacillus actinomycetemcomitans.
 XX
 PN WO200111081-A2.
 PD 15-FEB-2001.
 XX
 PF 04-AUG-2000; 2000WO-US021340.
 PF 06-AUG-1999; 99US-0147551P.
 PR
 XX
 PA (IVIG-) IVIGENE CORP.
 PI Prognulske-Fox A, Handfield M, Brady LJ, Hillman JD;
 PI Prognulske-Fox A, Handfield M, Brady LJ, Hillman JD;
 XX N-PSDB; AAF59422.
 DR WPI: 2001-202779/20.
 DR
 XX
 PT Identifying microbial polynucleotides, useful for vaccine design,
 PT diagnostics and antibiotherapy, comprises isolating clones of a microbe's
 PT expression library reactive with antibodies against microbe proteins
 PT produced during in vivo growth.
 XX
 PS Example 3; Page 65; 68pp; English.
 CC
 CC The present invention describes a method (M1) for identifying a
 CC polynucleotide (I) of a microbe (M) that is expressed in vivo. (M1)
 CC comprises: (a) absorbing antibodies (Ab) against antigens that are
 CC expressed by (M) in vivo and in vitro with cells or cellular extracts of
 CC (M) that have been grown in vitro; (b) isolating unadsorbed Abs; and (c)
 CC probing an expression library of (M)'s DNA/RNA with unadsorbed Abs, where
 CC (I) that is expressed in vivo is identified. The method can be used for
 CC identifying antigens expressed during an actual microbial infection. The
 CC identified polynucleotides are useful for vaccine design, diagnostics and
 CC antibiotherapy, in particular for the diagnosis and therapy of
 CC Actinobacillus actinomycetemcomitans infection, which is the etiologic
 CC agent for localised juvenile periodontitis. The present sequence

CC represents an *Actinobacillus actinomycetemcomitans* immunoreactive
 CC protein, which is used in an example from the present invention
 CC
 XX
 SQ Sequence 67 AA:

Query Match 4.0%; Score 11; DB 4; Length 67;
 Best Local Similarity 100.0%; Pred. No. 0.027; 0; Indels 0; Gaps 0;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 TIKGVVAGPE 48
 |||||
 DB 19 TIKGVVAGPE 29

RESULT 15

ABU27902
 ID ABU27902 standard; protein; 271 AA.

AC ABU27902;

DT 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #13429.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Enterobacter cloacae.

XX WO200277183-A2.

XX 03-OCT-2002.

PF 21-MAR-2002; 2002WO-US009107.

PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

DR N-PSDB; ACAA31772.

XX New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 55826; 1766bp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent

CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pcr_sequences

SQ Sequence 271 AA;

Query Match 4.0%; Score 11; DB 6; Length 271;
 Best Local Similarity 100.0%; Pred. No. 0.088; 0; Indels 0; Gaps 0;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 224 FVEDKDSPPYN 234
 |||||
 DB 219 FVEDKDSPPYN 229

RESULT 16

ABU48271
 ID ABU48271 standard; protein; 271 AA.

AC ABU48271;

DT 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #33798.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX *Salmonella typhi*.

XX WO200277183-A2.

XX 03-OCT-2002.

PF 21-MAR-2002; 2002WO-US009107.

PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

DR N-PSDB; ACAA52141.

XX New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 76195; 1766bp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 271 AA;

Query Match Best Local Similarity 4.0%; Score 11; DB 6; Length 271;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 224 FVEDKSDSPYVN 234
|||||
DB 219 FVEDKSDSPYVN 229

RESULT 17

ABU16868 standard; protein; 276 AA.

AC ABU16868;

DT 19-JUN-2003 (first entry)

DE Protein encoded by prokaryotic essential gene #2395.

KM Antisense; prokaryotic essential gene; cell proliferation; drug design.

OS Acinetobacter baumannii.

PN WO200277183-A2.

PD 03-OCT-2002.

PF 21-MAR-2002; 2002MO-US009107.

PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362659P.

PA (ELIT-) ELITRA PHARM INC.

PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,

PI Wall D, Trawick JD, Carr GT, Yamamoto R, Forsyth RA, Xu HH,

DR N-PSDB; ACA20738.

XX WPI: 2003-029926/02.

XX New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 44792; 17666p; English.

XX The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 276 AA;

Query Match Best Local Similarity 3.6%; Score 10; DB 6; Length 276;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 195 ARAIDVDVLA 204
|||||
DB 190 ARAIDVDVLA 199

RESULT 18

ADA35415 standard; protein; 286 AA.

AC ADA35415;

DT 20-NOV-2003 (first entry)

DE Acinetobacter baumannii protein #2576.

KM Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;

OS plant biocontrol agent.

PN Acinetobacter baumannii.

PN US6562958-B1.

PD 13-MAY-2003.

PF 04-JUN-1999; 99US-00328352.

PR 09-JUN-1998; 98US-0088701P.

PA (GENO-) GENOME THERAPEUTICS CORP.

PI Breton G, Bush D;

DR N-PSDB; ADA31289.

XX New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
 PT for diagnosing a bacterial disease, as components of antibacterial
 PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
 PT plants.
 XX
 XX Example; SEQ ID NO 6702; 328bp; English.
 CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
 CC The A. baumannii nucleic acids and polypeptides are useful as reagents
 CC for diagnosing a bacterial disease, as components of antibacterial
 CC vaccines, as targets for antibacterial drugs, to detect the presence of
 CC A. baumannii and other Acinetobacter species in a sample, in screening
 CC compounds for the ability to interfere with the A. baumannii life cycle
 CC or to inhibit A. baumannii infection, and as biocontrol agents for
 CC plants. The present sequence represents the amino acid sequence of an A.
 CC baumannii protein.
 CC
 SQ Sequence 286 AA;
 Query Match 3.6%; Score 10; DB 6; Length 286;
 Best Local Similarity 100.0%; Pred. No. 0.8;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 195 ARAIDVDVDA 204
 Db 200 ARAIDVDVDA 209
 RESULT 19
 ABR53692
 ID ABR53692 standard; protein; 125 AA.
 AC ABR53692;
 XX 20-JUN-2003 (first entry)
 DT Protein sequence #SEQ ID 2249.
 DE
 XX
 XX Multi-protein complex; eukaryote; drug target; diagnosis.
 KM
 XX Saccharomyces cerevisiae.
 OS
 XX
 XX EPI258494-A1.
 PN
 XX 20-NOV-2002.
 PD
 XX 20-DEC-2001; 2001EP-00130253.
 PF
 XX 15-MAY-2001; 2001EP-00111774.
 PR
 XX (CELL-) CELLZONE AG.
 PA
 XX Bauer A, Gavin A, Grandi P, Krause R, Kruse UD, Kuester BD;
 PI Marzloch M, Schultz JD, Superti-Furga GD;
 PI
 XX WPI: 2003-250078/25.
 DR N-PSDB; ACC61734.
 DR
 XX
 XX New isolated protein complexes useful for diagnosing a disease or
 PT disorder, or as a target for an active agent of a pharmaceutical,
 PT preferably a drug target in the treatment or prevention of disease or
 PT disorder.
 PS Disclosure; SEQ ID NO 2249; 17pp + Sequence listing; English.
 CC The invention relates to multiprotein complexes from eukaryotes. Proteins
 CC of the invention and DNA sequences encoding them are given in records
 CC ABR53692-ABR53903 and ACC60610-ACC61944 respectively. The complexes are
 CC obtainable by using a protein as a bait and isolating the set of proteins
 CC which is attached thereto from cells. Such protein complexes may comprise
 CC up to 30 distinct proteins. Protein complexes of the invention are useful
 CC for diagnosing a disease or disorder, or as a target for an active agent

CC of a pharmaceutical, preferably a drug target in the treatment or
 CC prevention of a disease or disorder. Note: The sequence data for this
 CC patent is not represented in the printed specification, but is based on
 CC sequence information supplied by the European Patent Office. The complete
 CC document is available on CD-ROM
 CC
 SQ Sequence 125 AA;
 Query Match 3.3%; Score 9; DB 6; Length 125;
 Best Local Similarity 100.0%; Pred. No. 3.4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 132 NELKDGATI 140
 Db 74 NELKDGATI 82
 RESULT 20
 ABU16864
 ID ABU16864 standard; protein; 241 AA.
 AC ABU16864;
 XX
 XX 19-JUN-2003 (first entry)
 DT Protein encoded by Prokaryotic essential gene #2391.
 DE
 XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 XX Acinetobacter baumannii.
 OS
 XX
 XX WO200277183-A2.
 PN
 XX 03-OCT-2002.
 PD
 XX 21-MAR-2002; 2002WO-US009107.
 PF
 XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 XX (ELIT-) ELITRA PHARM INC.
 PA
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 PI
 XX WPI: 2003-029926/02.
 DR N-PSDB; ACR20734.
 DR
 XX
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 PT
 XX Claim 25; SEQ ID NO 44788; 1766pp; English.
 PS
 XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an

organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp://wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences)

polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-regulated gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp://wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences)

Query Match 3.3%; Score 9; DB 6; Length 241;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Sequence 256 AA;

ABU33203
ID ABU33203 standard; protein; 256 AA.

ABU33203;

19-JUN-2003 (first entry)

Protein encoded by Prokaryotic essential gene #18730.

Antisense; prokaryotic essential gene; cell proliferation; drug design.

Legionella pneumophila.

MO200277183-A2.

03-OCT-2002.

21-MAR-2002; 2002WC-US009107.

21-MAR-2001; 2001US-00815242.

06-SEP-2001; 2001US-00948993.

25-OCT-2001; 2001US-0342923P.

08-FEB-2002; 2002US-00072851.

06-MAR-2002; 2002US-0362699P.

(ELIT-) ELITRA PHARM INC.

Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

WPI; 2003-029926/02.

N-PSDB; ACA37073.

New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.

Claim 25; SEQ ID NO 61127; 1766bp; English.

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated

Query Match 3.3%; Score 9; DB 6; Length 256;
Best Local Similarity 100.0%; Pred. No. 6.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 IAVPNDPSN 148
DB 125 IAVPNDPSN 133

RESULT 22

ABU23447
ID ABU23447 standard; protein; 262 AA.

ABU23447;

19-JUN-2003 (first entry)

Protein encoded by Prokaryotic essential gene #8974.

Antisense; prokaryotic essential gene; cell proliferation; drug design.

Bordetella pertussis.

MO200277183-A2.

03-OCT-2002.

21-MAR-2002; 2002WC-US009107.

21-MAR-2001; 2001US-00815242.

06-SEP-2001; 2001US-00948993.

25-OCT-2001; 2001US-0342923P.

08-FEB-2002; 2002US-00072851.

06-MAR-2002; 2002US-0362699P.

(ELIT-) ELITRA PHARM INC.

Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

WPI; 2003-029926/02.

N-PSDB; ACA27317.

New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 51371; 1766bp; English.
 CC
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-regulated gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 CC
 SQ Sequence 262 AA;
 Query Match 3.3%; Score 9; DB 6; Length 262;
 Best Local Similarity 100.0%; Pred. No. 6.4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 133 ELKDGATTA 141
 |||||
 Db 119 ELKDGATTA 127
 RESULT 23
 ABU27880
 ID ABU27880 standard; protein; 270 AA.
 XX
 AC ABU27880;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #13407.
 XX
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 OS Enterobacter cloacae.
 XX
 PN WO200277183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WC-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX

PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GT, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR MPI: 2003-029926/02.
 DR N-Psdb; AC31750.
 XX
 PS Claim 25; SEQ ID NO 55804; 1766bp; English.
 CC
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-regulated gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 CC
 SQ Sequence 270 AA;
 Query Match 3.3%; Score 9; DB 6; Length 270;
 Best Local Similarity 100.0%; Pred. No. 6.6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 176 DIVENPKKL 184
 |||||
 Db 165 DIVENPKKL 173
 RESULT 24
 ABM67576
 ID ABM67576 standard; protein; 272 AA.
 XX
 AC ABM67576;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Photorhabdus luminescens protein sequence #673.
 XX
 KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
 KW detection; food; gene expression; plant; animal; microorganism; toxin;
 KW antibiotic; biopesticide; virulence factor; disease model; plague;
 KW whooping cough.
 XX
 OS Photorhabdus luminescens.

XX MO200294867-A2.
XX 28-NOV-2002.
XX 07-FEB-2002; 2002MO-IB003040.
XX 07-FEB-2001; 2001FR-00001659.
XX (INSP) INST PASTEUR.
XX (CNRS) CNRS CENT NAT RECH SCI.
XX Duchaud E, Taourit S, Glaeser P, Frangeul L, Kunst F, Danchin A,
XX Buchrieser C,
XX WPI; 2003-148459/14.
XX
XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
XX useful e.g. as therapeutic antimicrobials and agricultural pesticides.
XX
XX Claim 2; SEQ ID NO 673; 1205bp; French.
XX
XX The invention relates to the isolation of genes and their encoded
XX proteins from Photorhabdus luminescens. The isolated sequences are
XX sources of probes and primers for detecting the genome of P. luminescens
XX and related species; to study polymorphisms; for gene analysis and for
XX detection/amplification of the genes. Antibodies (Ab) raised against the
XX polypeptides encoded by the genes are used for detection/identification
XX of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
XX carry a gene-containing vector are used to select compounds that
XX modulate, regulate, induce or inhibit expression of the genes in plants,
XX animals or microorganisms other than P. luminescens and are able to alter
XX response or sensitivity to toxins and antibiotics produced by P.
XX recombinant production of the proteins, particularly toxins and
XX antibacterials useful as insecticides, bactericides and fungicides. The
XX genes, proteins, vectors containing the genes and Ab are also useful
XX therapeutically (to treat microbial infection by bacteria or fungi that
XX are sensitive to P. luminescens-encoded toxins or antibiotics) and as
XX biopesticides. Other uses of the genes and the proteins are as vaccine
XX factors and for identifying targets of human diseases for which P.
XX luminescens is a model (particularly plague and whooping cough). This
XX sequence represents one of the isolated P. luminescens proteins
XX
XX Sequence 272 AA;
SQ
Query Match 3.3%; Score 9; DB 6; Length 272;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 158 KQGLIKLKD 166
DB 151 KQGLIKLKD 159
RESULT 25
ABBA48120
ID ABBA48120 standard; protein; 273 AA.
XX
XX ABB48120;
XX
XX 05-FEB-2002 (first entry)
XX
XX Listeria monocytogenes protein #824.
XX
XX Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
XX vitamin B12; bacterial infection; disease.
XX
XX Listeria monocytogenes.
XX
XX WO200177335-A2.
XX
XX 18-OCT-2001.
PD

XX 11-APR-2001; 2001MO-FR001118.
XX 11-APR-2000; 2000FR-00004629.
XX (INSP) INST PASTEUR.
XX
XX Buchrieser C, Frangeul L, Couve E, Ruenick C, Faini H, Dehoux P,
XX Dussurget O, Chetoui F, Nedjati H, Glaeser P, Kunst F, Cossart P,
XX Daniels U, Goebel W, Kreft U, Kuhn M, Ng E, Vazquez-Boland JA,
XX Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A,
XX Chakraborty T, Dommann E, Hain T, Berche P, Charbit A, Durant L,
XX Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N,
XX Madueno E, De Pablos B, Wehland U, Kaerst U, Entian K, Hauf J,
XX Rose M, Voss H;
XX WPI; 2002-010914/01.
XX
XX Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
XX and prevention of Listeria and related bacterial infections, and related
XX polypeptides.
XX
XX Claim 6; SEQ ID NO 825; 192bp; French.
XX
XX The present invention relates to the genome sequence of Listeria
XX monocytogenes EGD-e (see AB03041). The genome sequence and fragments of
XX it are useful for selecting probes and primers for detecting genes in L.
XX monocytogenes and related organisms, and for studying genetic
XX polymorphisms and other genomes. The present sequence is a protein
XX encoded from the genome sequence of the present invention. Proteins
XX expressed from the genome sequence are useful for raising specific
XX antibodies. Identification of L. monocytogenes and related organisms, and
XX for biosynthesis and biodegradation, especially biosynthesis of Vitamin
XX B12. The genome sequence and proteins encoded by it are also useful for
XX selecting compounds that regulate gene expression and cell replication
XX and modulate L. monocytogenes-related diseases. In addition, the genome
XX sequence and proteins encoded by it are useful in pharmaceutical and
XX vaccine compositions for the treatment or prevention of infections by L.
XX monocytogenes and related organisms. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIP0 at
XX ftp.wipo.int/pub/published_pcc_sequences
XX
XX Sequence 273 AA;
SQ
Query Match 3.3%; Score 9; DB 5; Length 273;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 140 IAVPNDPSN 148
DB 136 IAVPNDPSN 144
RESULT 26
ABU32634
ID ABU32634 standard; protein; 273 AA.
XX
XX ABU32634;
XX
XX 19-JUN-2003 (first entry)
XX
XX Protein encoded by Prokaryotic essential gene #18161.
XX
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
XX Listeria monocytogenes.
XX
XX WO200277183-A2.
XX
XX 03-OCT-2002.
XX
XX 21-MAR-2002; 2002MO-US009107.
PF

RESULT 28
 AC97607
 ID ADC97607 standard; protein: 291 AA.
 XX
 AC ADC97607;
 XX
 DT 01-JAN-2004 (first entry)
 XX
 DE E. faecium protein sequence SEQ ID 7234.
 XX
 KW Vaccine; urinary tract infection; bacteraemia; endocarditis; wound;
 KM abdominal-pelvic infection.
 XX
 OS Enterococcus faecium.
 XX
 PN US6583275-B1.
 XX
 PD 24-JUN-2003.
 XX
 PF 30-JUN-1998; 98US-00107532.
 XX
 PR 02-JUL-1997; 97US-0051571P.
 XX
 PR 14-MAY-1998; 98US-0085598P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Doucette-Stamm LA, Bush D;
 XX
 DR WPI, 2003-799836/75.
 XX
 DR N-PSDB; ADC93953.
 XX
 PT New isolated nucleic acid derived from Enterococcus faecium encoding an
 PT Enterococcus faecium polypeptide useful for detection, prevention and
 PT treatment of a pathological condition resulting from a bacterial
 PT infection.
 XX
 PS Example 1; SEQ ID NO 7234; 243bp; English.
 XX
 CC The invention relates to an isolated nucleic acid derived from
 CC Enterococcus faecium encoding an Enterococcus faecium polypeptide having
 CC one of 10 fully defined sequences given in the (or comprising) 40
 CC sequential nucleotides chosen from any of the nucleic acids, its
 CC complement or sequences hybridizing to it). Also included are a
 CC recombinant vector comprising the nucleic acid operably linked to
 CC transcription regulatory element, a cell comprising the vector and a
 CC single-stranded probe comprising the nucleic acid. The nucleic acids are
 CC chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.
 CC The nucleic acid is useful for diagnosing pathological conditions
 CC resulting from E. faecium bacterial infection (e.g. urinary tract
 CC infection, bacteraemia, endocarditis, wounds and abdominal-pelvic
 CC infection) and for screening drugs such as agonists and antagonists. The
 CC nucleic acid is useful for recombinant production of Candida albicans -
 CC derived peptides or antisense polypeptides. Pharmaceutical compositions
 CC and vaccines containing the nucleic acid are useful for preventing or
 CC treating Enterococcus faecium infections. The present sequence represents
 CC one if the disclosed E. faecium proteins.
 XX
 SQ Sequence 291 AA;
 XX
 Query Match 3.3%; Score 9; DB 7; Length 291;
 Best Local Similarity 100.0%; Pred. No. 7;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 231 PYVNIVAR 239
 DB 240 PYVNIVAR 248
 XX
 RESULT 29
 ADA35637
 ID ADA35637 standard; protein: 296 AA.
 XX

AC ADA35637;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Acinetobacter baumannii protein #2798.
 XX
 KW Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
 KM plant biocontrol agent.
 XX
 OS Acinetobacter baumannii.
 XX
 PN US6562958-B1.
 XX
 PD 13-MAY-2003.
 XX
 PF 04-JUN-1999; 99US-00328352.
 XX
 PR 09-JUN-1998; 98US-0088701P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Breton G, Bush D;
 XX
 DR WPI, 2003-576092/54.
 XX
 DR N-PSDB; ADA31511.
 XX
 PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
 PT for diagnosing a bacterial disease, as components of antibacterial
 PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
 PT plants.
 XX
 PS Example; SEQ ID NO 6924; 328bp; English.
 XX
 CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
 CC The A. baumannii nucleic acids and polypeptides are useful as reagents
 CC for diagnosing a bacterial disease, as components of antibacterial
 CC vaccines, as targets for antibacterial drugs, to detect the presence of
 CC A. baumannii and other Acinetobacter species in a sample, in screening
 CC compounds for the ability to interfere with the A. baumannii life cycle
 CC or to inhibit A. baumannii infection, and as biocontrol agents for
 CC plants. The present sequence represents the amino acid sequence of an A.
 CC baumannii protein.
 XX
 SQ Sequence 296 AA;
 XX
 Query Match 3.3%; Score 9; DB 6; Length 296;
 Best Local Similarity 100.0%; Pred. No. 7.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 148 NLARALLIL 156
 DB 164 NLARALLIL 172
 XX
 RESULT 30
 ID AAU18017
 ID AAU18017 standard; protein: 96 AA.
 XX
 AC AAU18017;
 XX
 DT 07-NOV-2001 (first entry)
 XX
 DE Human immunoglobulin polypeptide SEQ ID NO 162.
 XX
 KW Immunoglobulin; signal transduction pathway protein; cancer;
 KM antisense therapy; gene therapy; neurological disorder; renal disorder;
 KM cardiovascular disorder; gastrointestinal disorder; pulmonary disorder;
 KM reproductive disorder; immune system disorder; proliferative disorder;
 KM muscular disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200155315-A2.

XX 02-AUG-2001. 2001WO-US0013326.
XX 17-JAN-2001. 2001WO-US0013326.
XX 31-JAN-2000. 2000US-0179065P.
PR 04-FEB-2000. 2000US-0180628P.
PR 24-FEB-2000. 2000US-0184664P.
PR 02-MAR-2000. 2000US-0186350P.
PR 16-MAR-2000. 2000US-0189874P.
PR 17-MAR-2000. 2000US-0190076P.
PR 18-APR-2000. 2000US-0198123P.
PR 19-MAY-2000. 2000US-0205515P.
PR 07-JUN-2000. 2000US-0209467P.
PR 28-JUN-2000. 2000US-0214886P.
PR 30-JUN-2000. 2000US-0215135P.
PR 07-JUL-2000. 2000US-0216647P.
PR 07-JUL-2000. 2000US-0216880P.
PR 11-JUL-2000. 2000US-0217487P.
PR 14-JUL-2000. 2000US-0217496P.
PR 14-JUL-2000. 2000US-0218290P.
PR 26-JUL-2000. 2000US-0220963P.
PR 14-AUG-2000. 2000US-0220964P.
PR 14-AUG-2000. 2000US-0224516P.
PR 14-AUG-2000. 2000US-0224519P.
PR 14-AUG-2000. 2000US-0225213P.
PR 14-AUG-2000. 2000US-0225214P.
PR 14-AUG-2000. 2000US-0225266P.
PR 14-AUG-2000. 2000US-0225267P.
PR 14-AUG-2000. 2000US-0225268P.
PR 14-AUG-2000. 2000US-0225270P.
PR 14-AUG-2000. 2000US-0225447P.
PR 14-AUG-2000. 2000US-0225757P.
PR 14-AUG-2000. 2000US-0225758P.
PR 14-AUG-2000. 2000US-0225759P.
PR 18-AUG-2000. 2000US-0226279P.
PR 22-AUG-2000. 2000US-0226881P.
PR 22-AUG-2000. 2000US-0226886P.
PR 22-AUG-2000. 2000US-0227182P.
PR 23-AUG-2000. 2000US-0227009P.
PR 30-AUG-2000. 2000US-0228924P.
PR 01-SEP-2000. 2000US-0229287P.
PR 01-SEP-2000. 2000US-0229343P.
PR 01-SEP-2000. 2000US-0229344P.
PR 01-SEP-2000. 2000US-0229345P.
PR 05-SEP-2000. 2000US-0229309P.
PR 05-SEP-2000. 2000US-0229513P.
PR 06-SEP-2000. 2000US-0230437P.
PR 06-SEP-2000. 2000US-0231242P.
PR 08-SEP-2000. 2000US-0231243P.
PR 08-SEP-2000. 2000US-0231244P.
PR 08-SEP-2000. 2000US-0231413P.
PR 08-SEP-2000. 2000US-0231414P.
PR 08-SEP-2000. 2000US-0232080P.
PR 08-SEP-2000. 2000US-0232081P.
PR 12-SEP-2000. 2000US-0233968P.
PR 14-SEP-2000. 2000US-0233397P.
PR 14-SEP-2000. 2000US-0233398P.
PR 14-SEP-2000. 2000US-0233399P.
PR 14-SEP-2000. 2000US-0233400P.
PR 14-SEP-2000. 2000US-0233401P.
PR 14-SEP-2000. 2000US-0233063P.
PR 14-SEP-2000. 2000US-0233064P.
PR 14-SEP-2000. 2000US-0233065P.
PR 21-SEP-2000. 2000US-0234223P.
PR 21-SEP-2000. 2000US-0234274P.
PR 25-SEP-2000. 2000US-0234997P.
PR 25-SEP-2000. 2000US-0234998P.
PR 26-SEP-2000. 2000US-0235484P.
PR 27-SEP-2000. 2000US-0235834P.
PR 27-SEP-2000. 2000US-0235836P.
PR 29-SEP-2000. 2000US-0236327P.
PR 29-SEP-2000. 2000US-0236367P.
PR 29-SEP-2000. 2000US-0236368P.
PR 29-SEP-2000. 2000US-0236369P.
PR 29-SEP-2000. 2000US-0236370P.
PR 02-OCT-2000. 2000US-0236802P.
PR 02-OCT-2000. 2000US-0237037P.
PR 02-OCT-2000. 2000US-0237038P.
PR 02-OCT-2000. 2000US-0237039P.
PR 02-OCT-2000. 2000US-0237040P.
PR 13-OCT-2000. 2000US-0239335P.
PR 13-OCT-2000. 2000US-0239337P.
PR 20-OCT-2000. 2000US-0240960P.
PR 20-OCT-2000. 2000US-0241221P.
PR 20-OCT-2000. 2000US-0241785P.
PR 20-OCT-2000. 2000US-0241786P.
PR 20-OCT-2000. 2000US-0241787P.
PR 20-OCT-2000. 2000US-0241808P.
PR 20-OCT-2000. 2000US-0241809P.
PR 20-OCT-2000. 2000US-0241826P.
PR 01-NOV-2000. 2000US-0244617P.
PR 08-NOV-2000. 2000US-0246474P.
PR 08-NOV-2000. 2000US-0246475P.
PR 08-NOV-2000. 2000US-0246476P.
PR 08-NOV-2000. 2000US-0246477P.
PR 08-NOV-2000. 2000US-0246478P.
PR 08-NOV-2000. 2000US-0246479P.
PR 08-NOV-2000. 2000US-0246523P.
PR 08-NOV-2000. 2000US-0246524P.
PR 08-NOV-2000. 2000US-0246525P.
PR 08-NOV-2000. 2000US-0246526P.
PR 08-NOV-2000. 2000US-0246527P.
PR 08-NOV-2000. 2000US-0246528P.
PR 08-NOV-2000. 2000US-0246529P.
PR 08-NOV-2000. 2000US-0246532P.
PR 08-NOV-2000. 2000US-0246609P.
PR 08-NOV-2000. 2000US-0246610P.
PR 08-NOV-2000. 2000US-0246611P.
PR 08-NOV-2000. 2000US-0246613P.
PR 17-NOV-2000. 2000US-0249207P.
PR 17-NOV-2000. 2000US-0249208P.
PR 17-NOV-2000. 2000US-0249209P.
PR 17-NOV-2000. 2000US-0249210P.
PR 17-NOV-2000. 2000US-0249211P.
PR 17-NOV-2000. 2000US-0249212P.
PR 17-NOV-2000. 2000US-0249213P.
PR 17-NOV-2000. 2000US-0249214P.
PR 17-NOV-2000. 2000US-0249215P.
PR 17-NOV-2000. 2000US-0249216P.
PR 17-NOV-2000. 2000US-0249217P.
PR 17-NOV-2000. 2000US-0249218P.
PR 17-NOV-2000. 2000US-0249244P.
PR 17-NOV-2000. 2000US-0249245P.
PR 17-NOV-2000. 2000US-0249246P.
PR 17-NOV-2000. 2000US-0249247P.
PR 17-NOV-2000. 2000US-0249297P.
PR 17-NOV-2000. 2000US-0249298P.
PR 17-NOV-2000. 2000US-0249300P.
PR 01-DEC-2000. 2000US-0250160P.
PR 01-DEC-2000. 2000US-0250391P.
PR 05-DEC-2000. 2000US-0251030P.
PR 05-DEC-2000. 2000US-0251988P.
PR 05-DEC-2000. 2000US-0251799P.
PR 06-DEC-2000. 2000US-0251479P.
PR 08-DEC-2000. 2000US-0251855P.
PR 08-DEC-2000. 2000US-0251856P.
PR 08-DEC-2000. 2000US-0251857P.
PR 08-DEC-2000. 2000US-0251858P.
PR 08-DEC-2000. 2000US-0251989P.
PR 08-DEC-2000. 2000US-0251990P.
PR 11-DEC-2000. 2000US-0254097P.
PR 05-JAN-2001. 2001US-0259678P.
(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;

DR MPI: 2001-45725/49.
DR N-PSDB; AAS28805.
XX Isolated novel immunoglobulin polypeptide for monitoring the presence and
PT progression of diseases and for diagnosis.
XX Claim 11; SEQ ID NO 162; 551pp; English.
XX Sequences AAU17977-AAU18087 represent immunoglobulin polypeptides of the
CC invention. The polypeptides and their associated polynucleotides can be
CC used to diagnose a pathological condition or a susceptibility to a
CC pathological condition in a subject by determining the presence or
CC absence of a mutation in a DNA sequence or determining the presence or
CC amount of expression of the protein. Alternatively the identification of
CC a binding partner to a sequence allows determination of changes in
CC protein activity. The sequences can be used as research tools for
CC receptors or other signal transduction pathway proteins that interact
CC with the polypeptides of the invention and can be used to treat, prevent
CC or diagnose various types of disorders such as neurological disorders,
CC cardiovascular disorders, gastrointestinal disorders, reproductive
CC disorders, immune system disorders, renal disorders, muscular disorders,
CC pulmonary disorders, proliferative disorders and cancer. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcc_sequences
XX
SQ Sequence 96 AA;
Query Match 2.9%; Score 8; DB 4; Length 96;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 146 PSNLRAL 153
Db 76 PSNLRAL 83
RESULT 31
ABBI0497
ID ABBI0497 standard; protein, 96 AA.
XX
AC ABBI0497;
XX
DT 10-JAN-2002 (first entry)
XX
DE Human CDNA SEQ ID NO: 805.
XX
XX Human; gene therapy; neural disorder; immune system disorder;
KM muscular disorder; reproductive disorder; gastrointestinal disorder;
KM pulmonary disorder; cardiovascular disorder; renal disorder;
KM proliferative disorder; inflammation.
XX
OS Homo sapiens.
XX
PV MO200154474-A2.
XX
FD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001349.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0160628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216890P.

PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0228928P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 06-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236369P.
PR 02-OCT-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.

PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251836P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX
XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
XX Rosen CA, Ruben SM, Barash SC;
PI WPI; 2002-681727/73.
XX N-PSDB; ABV84056.
DR
XX
XX Novel polypeptide useful for diagnosis, prognosis, prevention, and
PT treatment of immune, hyperproliferative, renal, respiratory,
PT cardiovascular, reproductive, endocrine, gastrointestinal and
PT neurological disorders.
XX
XX
XX Claim 11; SEQ ID NO 805; 369pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (ABV83682-ABV84101) and proteins
CC (ABP6710-ABP67129) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breasts, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 96 AA;
SQ
Query Match 2.9%; Score 8; DB 5; Length 96;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 146 PSNLARAL 153
Db 76 PSNLARAL 83
RESULT 33
ADBJ1641
ID ADBJ1641 standard; protein; 96 AA.
XX
XX ADBJ1641;
AC
XX
DT 04-DEC-2003 (first entry)
XX
XX Human novel protein SEQ ID NO 162.
XX
XX Gene therapy; human; immunoglobulin; cancer; lung cancer; leukaemia;
KM ovarian cancer; epithelial cancer; hyperplasia; Gaucher's disease; AIDS;
KM arrhythmia; cardiac oedema; ischaemia; pneumonia; cystic fibrosis;
KM asthma; sarcoidosis; rhinitis; anaemia; inflammation; sinusitis;
XX chronic obstructive pulmonary disease; infectious disease.
XX

OS Homo sapiens.
XX
XX US2003077606-A1.
XX
XX 24-APR-2003.
XX
XX 07-MAR-2002; 2002US-00091438.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226881P.
PR 22-AUG-2000; 2000US-0226886P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.

PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236357P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239393P.
 PR 13-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249246P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 06-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251983P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 PR 17-JAN-2001; 2001US-0076487P.
 XX

PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM, Barash SC;
 XX WPI; 2003-615993/58.
 DR N-PSDB; ADB31530.
 XX
 XX New human immunoglobulin superfamily of polypeptide and genes, useful for
 PT treating, preventing or diagnosing e.g. cancers (lung cancer, leukemia,
 PT ovarian epithelial cancer, etc.), hyperplasia, ischemia, pneumonia or
 PT AIDS.
 XX
 PS Claim 12; SEQ ID NO 162; 213pp; English.
 XX
 XX The invention relates to an isolated polypeptide, which comprises the
 CC human immunoglobulin superfamily of proteins. The polypeptide or
 CC polynucleotide is useful for treating, preventing or ameliorating a
 CC medical condition e.g. cancers (lung cancer, leukemia, ovarian
 CC epithelial cancer, etc.), hyperplasia, Gaucher's disease, AIDS,
 CC arrhythmia, cardiac oedema, ischaemia, pneumonia, cystic fibrosis,
 CC asthma, sarcoidosis, rhinitis, anaemia, leukaemia, inflammatory
 CC sinusitis, chronic obstructive pulmonary disease, infectious diseases
 CC etc. The polypeptide or polynucleotide is also useful for diagnosing any
 CC of these diseases or a susceptibility to the disease. The present
 CC sequence represent the amino acid sequence of a novel human protein.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format direct from USPTO at
 CC seqdata.uspto.gov/sequence.html?docid=20030077606
 XX
 SQ Sequence 96 AA;
 XX
 QY 146 PSNLARAL 153
 DB 76 PSNLARAL 83
 XX
 RESULT 34
 ID AAO05643 standard; protein; 152 AA.
 AC AAO05643;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 19535.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation.
 XX
 OS Homo sapiens.
 XX
 PN WO200164835-A2.
 XX
 PD 07-SEP-2001.
 XX
 PF 26-FEB-2001; 2001WO-US004927.
 XX
 PR 28-FEB-2000; 2000US-00515126.
 PR 18-MAY-2000; 2000US-00577409.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX WPI; 2001-514838/56.
 DR N-PSDB; AAI85574.
 XX

PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
PS Claim 20; SEQ ID NO 19535; 1339pp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AA179941-AA193841) and
CC the encoded proteins (AA000010-AA013910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukemia, nervous system disorders, arthritis and
CC inflammation. Notes: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at http://wipo.int/pub/published_pct_sequences
SQ Sequence 152 AA;

Query Match 2.9%; Score 8; DB 4; Length 152;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 PSNLABAL 153
DB 116 PSNLABAL 123

RESULT 35

AAU36063 ID AAU36063 standard; protein; 240 AA.

AC AAU36063;

DT 14-FEB-2002 (first entry)

DE Klebsiella pneumoniae cellular proliferation protein #51.

KW Antisense; prokaryotic cellular proliferation protein; antibiotic;

KW antibacterial; drug design.

OS Klebsiella pneumoniae.

PN MO200170955-A2.

PD 27-SEP-2001.

PF 21-MAR-2001; 2001WO-US009180.

PR 21-MAR-2000; 2000US-0191078P.

PR 23-MAY-2000; 2000US-0206848P.

PR 26-MAY-2000; 2000US-0207172P.

PR 23-OCT-2000; 2000US-0242578P.

PR 27-NOV-2000; 2000US-0253625P.

PR 22-DEC-2000; 2000US-0257931P.

PR 16-FEB-2001; 2001US-0269308P.

PA (ELIT-) ELITRA PHARM INC.

PI Haselbeck R, Ohlsen KL, Zykkind JW, Wall D, Trawick JD, Carr GT,

PI Yamamoto RT, Xu HH;

DR WPI; 2001-611495/70.

DR N-PSDB; AAS53922.

XX New polynucleotides for the identification and development of

XX antibiotics, comprise sequences of antisense nucleic acids.

XX Example 3; SEQ ID NO 11656; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to

XX

XX

XX

CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes
CC themselves and the encoded proteins. The prokaryotes used are Escherichia
CC coli; Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence represents an essential prokaryotic
CC cellular proliferation protein. Notes: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC http://wipo.int/pub/published_pct_sequences
SQ Sequence 240 AA;

Query Match 2.9%; Score 8; DB 4; Length 240;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 VGNTFYVP 120
DB 75 VGNTFYVP 82

RESULT 36

ABP26478 ID ABP26478 standard; protein; 263 AA.

AC ABP26478;

DT 02-JUL-2002 (first entry)

DE Streptococcus polypeptide SEQ ID NO 2132.

KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;

KW group A streptococcus; Streptococcus pyogenes; antibacterial;

KW antiinflammatory; infection; vaccine; meningitis; gene therapy.

OS Streptococcus pyogenes.

PN MO200234771-A2.

PD 02-MAY-2002.

PF 29-OCT-2001; 2001WO-GB004789.

PR 27-OCT-2000; 2000GB-00026333.

PR 24-NOV-2000; 2000GB-00028727.

PR 07-MAR-2001; 2001GB-00005640.

PA (CHIR-) CHIRON SPA.

PA (GENO-) INST GENOMIC RES.

PI Telford J, Massignani V, Margarit Y Rosi, Grandi G, Fraser C;

PI Tettelein H;

DR WPI; 2002-352536/38.

DR N-PSDB; AEN67109.

XX New Streptococcus protein for the treatment or prevention of infection or

XX disease caused by Streptococcus bacteria, such as meningitis, and for

XX detecting a compound that binds to the protein.

XX Claim 1; Page 3365; 4525pp; English.

XX The invention relates to a protein (ABP25413-ABP30895) from group B

XX Streptococcus/GAS (Streptococcus agalactiae) or group A streptococcus/GAS

XX

XX

XX

XX

the specification. The proteins have antibacterial and anti-inflammatory activity. (1), nucleic acids encoding (1), ABN65044-ABN71526 and CC antibodies that bind (1) are used in the manufacture of medicaments for the treatment or prevention of infection or disease caused by CC Streptococcus bacteria, particularly *S. agalactiae* and *S. pyogenes*. CC Nucleic acids encoding (1) are used to detect Streptococcus in a biological sample. (1) is used to determine whether a compound binds to CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be used as a vaccine or diagnostic composition. The disease caused by CC Streptococcus that is prevented or treated may be meningitis. Nucleic acid encoding (1) may be used to recombinantly produce (1) and may be used in gene therapy. Antibodies to (1) are used for affinity chromatography, immunoassays, and distinguishing/identifying CC Streptococcus proteins

XX
SQ Sequence 263 AA;

Query Match 2.9%; Score 8; DB 5; Length 263;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 KAYOTDEV 260
240 KAYOTDEV 247

RESULT 37
ABU25052
ID ABU25052 standard; protein; 263 AA.
XX
AC ABU25052;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #10579.
XX
KM Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Clostridium difficile.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002MO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
XX
DR N-PSDB; ACA28922.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 52976; 1766bp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the

antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular CC proliferation or the activity of a gene in an operon required for CC the gene product; (7) identifying a compound that influences the activity of CC the gene product or that has an activity against a biological pathway CC required for proliferation, or that inhibits cellular proliferation; (8) CC identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies CC or a gene on which the test compound that inhibits proliferation of an CC organism acts; (9) manufacturing an antibiotic; (10) profiling a CC compound's activity; (11) a culture comprising strains in which the gene CC product is overexpressed or underexpressed; (12) determining the extent CC to which each of the strains is present in a culture or collection of CC strains; or (13) identifying the target of a compound that inhibits the CC proliferation of an organism. The antisense nucleic acids are useful for CC identifying proteins or screening for homologous nucleic acids required CC for cellular proliferation to isolate candidate molecules for rational CC drug discovery programs, or for screening homologous nucleic acids CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*, CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of CC the target prokaryotic essential genes. Note: The sequence data for this CC patent did not form part of the printed specification, but was obtained CC in electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 263 AA;

Query Match 2.9%; Score 8; DB 6; Length 263;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 229 DSPYVNI 236
216 DSPYVNI 223

Db

RESULT 38
ABU23277
ID ABU23277 standard; protein; 265 AA.
XX
AC ABU23277;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #804.
XX
KM Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Bordetella pertussis.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002MO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
XX
DR N-PSDB; ACA27147.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 51201; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway of
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed; (12) determining the extent
XX to which each of the strains is present in a culture or collection of
XX strains; or (13) identifying the target of a compound that inhibits the
XX proliferation of an organism. The antisense nucleic acids are useful for
XX identifying proteins or screening for homologous nucleic acids required
XX for cellular proliferation to isolate candidate molecules for rational
XX drug discovery programs, or for screening homologous nucleic acids
XX required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
XX *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
XX the target prokaryotic essential genes. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences

XX Sequence 265 AA;

Query Match 2.9%; Score 8; DB 6; Length 265;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 131 LNELEKGA 138
Db 120 LNELEKGA 127

RESULT 39

ADC96816
ID ADC96816 standard; protein; 269 AA.

AC ADC96816;

DT 01-JAN-2004 (first entry)

DE E. faecium protein sequence SEQ ID 6443.

XX Vaccine; urinary tract infection; bacteraemia; endocarditis; wound;
XX abdominal-pelvic infection.

OS Enterococcus faecium.

PN US6583275-B1.

PD 24-JUN-2003.

PF 30-JUN-1998; 98US-00107532.

PR 02-JUL-1997; 97US-0051571P.

XX 14-MAY-1998; 98US-0085598P.

PA (GENO-) GENOME THERAPEUTICS CORP.
XX Doucette-Stamm LA, Bush D;
PI

XX WPI; 2003-799836/75.
DR N-PSDB; ADC93162.

XX New isolated nucleic acid derived from *Enterococcus faecium* encoding an
XX protein. The nucleic acid is useful for detection, prevention and
XX treatment of a pathological condition resulting from a bacterial
XX infection.

XX Example 1; SEQ ID NO 6443; 243pp; English.

XX The invention relates to an isolated nucleic acid derived from
XX *Enterococcus faecium* encoding an *Enterococcus faecium* polypeptide having
XX one of 10 fully defined sequences given in the (or comprising 40
XX sequential nucleotides chosen from any of the nucleic acids, its
XX complement or sequences hybridizing to it). Also included are a
XX recombinant vector comprising the nucleic acid operably linked to
XX transcription regulatory element, a cell comprising the vector and a
XX single-stranded probe comprising the nucleic acid. The nucleic acids are
XX chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.
XX The nucleic acids are useful for diagnosing pathological conditions
XX resulting from *E. faecium* bacterial infection (e.g. urinary tract
XX infection, bacteraemia, endocarditis, wounds and abdominal-pelvic
XX infection) and for screening drugs such as agonists and antagonists. The
XX nucleic acid is useful for recombinant production of Candida albicans -
XX derived peptides or antisense polypeptides. Pharmaceutical compositions
XX and vaccines containing the nucleic acid are useful for preventing or
XX treating *Enterococcus faecium* infections. The present sequence represents
XX one of the disclosed *E. faecium* proteins.

XX Sequence 269 AA;

Query Match 2.9%; Score 8; DB 7; Length 269;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 ELKDGATI 140
Db 118 ELKDGATI 125

RESULT 40

ABU21415
ID ABU21415 standard; protein; 270 AA.

AC ABU21415;

DT 19-JUN-2003 (first entry)

DE Protein encoded by prokaryotic essential gene #6942.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

OS Burkholderia fungorum.

PN WO200277183-A2.

PD 03-OCT-2002.

PF 21-MAR-2002; 2002MO-US009-107.

PR 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362689P.

PA (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KU, Zyskind JW;
XX Mail D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.

DR N-PSDB; ACA25285.

XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.

PS Claim 25; SEQ ID NO 45339; 1766bp; English.

XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 270 AA;

Query Match 2.9%; Score 8; DB 6; Length 270;

Best Local Similarity 100.0%; Pred. No. 57;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 141 AVENDPSN 148

Db 132 AVENDPSN 139

Search completed: June 16, 2004, 11:15:51
Job time : 90 secs

Tue Jun 22 11:04:03 2004

us-10-018-672-2.oligo.rpr

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 16, 2004, 11:13:13 ; Search time 21 Seconds

(Without alignments)
1264.231 Million cell updates/sec

Title: US-10-018-672-2

Sequence: 1 MNFGKINGICALAGALAG.....TDEVEAKKQKGVKGM 276

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283366 seqs, 96191526 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database : PIR 78:*

1: Dir1:*
2: Dir2:*
3: Dir3:*
4: Dir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	12.3	277	2 JN0751	Outer membrane 30K
2	19	6.9	273	2 B64082	outer membrane pro
3	12	4.3	263	2 JN0753	outer membrane 29.
4	11	4.0	271	2 G90653	probable lipoprote
5	11	4.0	271	2 G85504	probable lipoprote
6	11	4.0	271	2 A10532	probable lipoprote
7	11	4.0	271	2 A48107	DNA-directed RNA p
8	9	3.3	273	2 A11110	probable lipoprote
9	9	3.3	273	2 A82286	lipoprotein yacC V
10	9	3.3	278	2 A13551	ABC transporter su
11	9	3.3	391	1 MHUBT	ig mu heavy chain
12	8	2.9	74	2 E81284	hypothetical prote
13	8	2.9	89	2 H75053	molybdopterin con
14	8	2.9	259	2 A83108	outer membrane lip
15	8	2.9	259	2 G98178	probable periplasm
16	8	2.9	271	2 E64744	probable lipoprote
17	8	2.9	446	2 A83213	probable ATP-depen
18	8	2.5	87	2 C81400	hypothetical prote
19	7	2.5	104	2 B88029	protein P46F5.13 (
20	7	2.5	114	2 G90820	hypothetical prote
21	7	2.5	114	2 G90851	hypothetical prote
22	7	2.5	114	2 G90874	hypothetical prote
23	7	2.5	114	2 C90902	hypothetical prote
24	7	2.5	114	2 D90911	hypothetical prote
25	7	2.5	114	2 F90971	hypothetical prote
26	7	2.5	114	2 H91065	hypothetical prote
27	7	2.5	114	2 D85679	unknown protein en
28	7	2.5	115	2 F85819	unknown protein en
29	7	2.5	115	2 H85627	hypothetical prote

30	7	2.5	123	2 F85744	unknown protein en
31	7	2.5	124	2 H64608	hypothetical prote
32	7	2.5	126	2 A82559	integrational host f
33	7	2.5	137	2 A64312	hypothetical prote
34	7	2.5	141	2 UB0359	plasmacyte sprea
35	7	2.5	154	2 S74334	biotin carboxyl ca
36	7	2.5	156	2 I67751	dopamine receptor
37	7	2.5	188	2 C97438	hypothetical prote
38	7	2.5	189	2 AF2656	conserved hypothet
39	7	2.5	193	2 J02185	coat protein - app
40	7	2.5	196	2 AC3629	3-octaprenyl-4-hyd
41	7	2.5	207	2 B53801	chitin synthase (E
42	7	2.5	215	2 T03680	plasma membrane pr
43	7	2.5	232	2 T03678	mrR protein precu
44	7	2.5	233	1 C39142	hypothetical prote
45	7	2.5	234	2 E71905	hypothetical prote
46	7	2.5	239	2 A81301	probable anion-upt
47	7	2.5	241	2 D87547	3-oxodipate CoA-t
48	7	2.5	247	2 F81410	hydrogenase isoenz
49	7	2.5	259	2 H64211	guanylate kinase (
50	7	2.5	260	2 B82957	probable TonB-depe
51	7	2.5	261	2 G82750	outer membrane pro
52	7	2.5	262	2 F82170	peptide ABC transp
53	7	2.5	268	2 J01473	pancreatic elastas
54	7	2.5	268	2 S39589	peptide transport
55	7	2.5	268	2 E64877	peptide transport
56	7	2.5	268	2 C90862	hypothetical prote
57	7	2.5	268	2 F85756	hypothetical prote
58	7	2.5	268	2 AC0656	peptide transport
59	7	2.5	269	2 E64134	peptide transport
60	7	2.5	271	2 A71802	probable outer mem
61	7	2.5	271	2 D64715	outer membrane pro
62	7	2.5	271	2 AG0131	probable lipoprote
63	7	2.5	271	2 AH0287	peptide transport
64	7	2.5	273	2 AB1472	probable lipoprote
65	7	2.5	274	2 AH0160	probable exported
66	7	2.5	276	2 JN0752	outer membrane 30.
67	7	2.5	276	2 AC1746	conserved lipoprot
68	7	2.5	276	2 A11376	conserved lipoprot
69	7	2.5	286	2 F86664	outer membrane lip
70	7	2.5	286	2 G86664	outer membrane lip
71	7	2.5	297	2 F64470	sulfate permease (
72	7	2.5	298	1 XWBO	ADP,ATP carrier pr
73	7	2.5	298	1 A29132	ADP,ATP carrier pr
74	7	2.5	298	1 S03894	ADP,ATP carrier pr
75	7	2.5	298	2 S31814	ADP,ATP carrier pr
76	7	2.5	298	2 B43646	ADP,ATP carrier pr
77	7	2.5	298	2 A43118	conserved hypothet
78	7	2.5	300	2 T15206	hypothetical prote
79	7	2.5	300	2 B75619	probable cobalamu
80	7	2.5	300	2 AC2831	transcription regu
81	7	2.5	300	2 C97609	transcription regu
82	7	2.5	306	2 G86168	hypothetical 27.0K
83	7	2.5	307	2 F86703	conserved hypothet
84	7	2.5	311	2 A81915	hydroxymethylbilan
85	7	2.5	311	2 AB1188	hydroxymethylbilan
86	7	2.5	314	2 T46743	carbamate kinase (
87	7	2.5	334	2 T20728	hypothetical prote
88	7	2.5	341	2 F75334	3-oxoacyl-acyl car
89	7	2.5	344	2 G70458	hypothetical prote
90	7	2.5	347	2 G00006	haploglobin - biac
91	7	2.5	350	2 S39156	outer membrane pro
92	7	2.5	352	2 F64751	probable ABC-type
93	7	2.5	352	2 A85531	hypothetical prote
94	7	2.5	352	2 B90680	probable ABC trans
95	7	2.5	380	2 B82634	threonine dehydrat
96	7	2.5	386	2 S75400	hypothetical prote
97	7	2.5	397	2 B87343	conserved hypothet
98	7	2.5	400	2 C82503	hypothetical prote
99	7	2.5	404	2 T36254	probable valine-py
100	7	2.5	406	2 AD0833	glucose membrane
101	7	2.5	412	1 S74961	glucose dehydrogen
102	7	2.5	419	2 G70652	probable sets prot

103	7	2.5	420	2	E70914	probable lipo prot	176	6	2.2	79	2	H42505	K-ORF-B protein -
104	7	2.5	432	2	A12613	hypothetical prote	177	6	2.2	82	2	A99836	hypothetical prote
105	7	2.5	432	2	G97395	hypothetical prote	178	6	2.2	82	2	C56594	unknown protein en
106	7	2.5	434	2	G75339	conserved hypotet	179	6	2.2	84	2	A81622	ribosomal protein
107	7	2.5	454	2	D69066	amidase - Methanob	180	6	2.2	84	2	A81259	ribosomal protein
108	7	2.5	457	2	I51660	dopamine D1B recep	181	6	2.2	85	2	G81150	hypothetical prote
109	7	2.5	468	2	T21331	hypothetical prote	182	6	2.2	90	2	F91254	hypothetical prote
110	7	2.5	471	2	S48265	hypothetical prote	183	6	2.2	90	2	D70678	hypothetical prote
111	7	2.5	473	2	H72419	virulence factor M	184	6	2.2	90	2	B86095	hypothetical prote
112	7	2.5	475	2	A41271	dopamine receptor	185	6	2.2	90	2	F65209	hypothetical 10.5
113	7	2.5	477	1	DYHND5	dopamine receptor	186	6	2.2	92	2	AF1012	conserved hypotet
114	7	2.5	486	2	B55886	dopamine receptor	187	6	2.2	92	2	B75559	ribosomal protein
115	7	2.5	491	2	UC5312	UTP-hexose-1-phosp	188	6	2.2	92	2	A47130	alcohol dehydrogen
116	7	2.5	530	2	D82063	probable thiamin A	189	6	2.2	92	2	T20113	hypothetical prote
117	7	2.5	538	2	G72539	probable CTP synth	190	6	2.2	93	2	C99141	hypothetical prote
118	7	2.5	549	2	H96944	arginine degradati	191	6	2.2	95	1	H18P14	internal protein I
119	7	2.5	553	2	E89725	protein CO2C6.2 [i	192	6	2.2	96	2	S48006	IPI protein - phag
120	7	2.5	554	2	T18858	hypothetical prote	193	6	2.2	97	2	H69835	hypothetical HupF -
121	7	2.5	572	2	T37128	hypothetical prote	194	6	2.2	98	2	D41892	hypothetical prote
122	7	2.5	575	2	B64174	hypothetical prote	195	6	2.2	98	2	D83872	hypothetical prote
123	7	2.5	580	2	S76668	hypothetical prote	196	6	2.2	99	2	A83452	utase operon 23X
124	7	2.5	600	2	A45112	major paraflogella	197	6	2.2	99	2	A60926	probable lipoprote
125	7	2.5	610	2	T47725	hypothetical prote	198	6	2.2	100	2	F64255	ribosomal protein
126	7	2.5	610	2	T16194	hypothetical prote	199	6	2.2	102	2	A53308	hypothetical prote
127	7	2.5	611	2	F75095	probable asparagin	200	6	2.2	104	2	AC1159	hypothetical prote
128	7	2.5	616	2	H82130	proteinase IV VCI9	201	6	2.2	104	2	AC1518	hypothetical prote
129	7	2.5	632	2	C95129	DNA ligase, NAD-de	202	6	2.2	106	1	Z2BPT9	gene 49.2 protein
130	7	2.5	652	2	H97959	DNA ligase (NAD) (203	6	2.2	106	2	T17814	hypothetical prote
131	7	2.5	670	2	B70145	periplasmic protei	204	6	2.2	107	2	F53306	K1es protein - pla
132	7	2.5	691	2	C64548	outer membrane pro	205	6	2.2	109	2	D70701	hypothetical prote
133	7	2.5	694	2	E72272	conserved hypotet	206	6	2.2	109	2	AG3602	dhypodolipoamide d
134	7	2.5	696	2	E95938	probable ATP-bind	207	6	2.2	110	2	G84101	flagellar protein
135	7	2.5	783	1	A48984	DNA excision repai	208	6	2.2	111	2	T00688	hypothetical prote
136	7	2.5	852	2	H81906	probable lipo-pro	209	6	2.2	112	2	G69781	thioredoxin homolo
137	7	2.5	852	2	B81110	protein-PIT uridy	210	6	2.2	113	2	H75183	hypothetical prote
138	7	2.5	875	2	AE1270	DNA polymerase I	211	6	2.2	116	2	SS9099	ribosomal protein
139	7	2.5	889	2	UC6015	chitin synthase (E	212	6	2.2	117	2	A83640	hypothetical prote
140	7	2.5	905	2	AH2892	hypothetical prote	213	6	2.2	118	2	G70306	hypothetical prote
141	7	2.5	905	2	C97668	ABC transporter re	214	6	2.2	121	2	A71031	hypothetical prote
142	7	2.5	1036	2	A82357	probable multidrug	215	6	2.2	121	2	B81872	ribosomal protein
143	7	2.5	1073	2	S14032	kinesin-related pr	216	6	2.2	122	1	S33697	hypothetical prote
144	7	2.5	1095	2	E96744	probable oligopept	217	6	2.2	124	2	H89794	hypothetical prote
145	7	2.5	1118	2	H97298	subtilisin like pr	218	6	2.2	125	2	H47677	hypothetical prote
146	7	2.5	1130	2	T20288	hypothetical prote	219	6	2.2	126	2	AD2313	hypothetical prote
147	7	2.5	1239	2	G02750	DNA-directed DNA p	220	6	2.2	128	2	T14487	vary hypotetrical
148	7	2.5	1333	2	S38635	blastocyst polyprot	221	6	2.2	128	2	H97101	uncharacterized pr
149	7	2.5	1395	2	F97713	190K antigen precu	222	6	2.2	128	2	H97354	uncharacterized pr
150	7	2.5	1994	2	D86452	protein F6N18.13	223	6	2.2	128	2	AF3006	acetyltransferase
151	7	2.5	2154	2	F83068	hypothetical prote	224	6	2.2	128	2	F82777	ketanuglycin acetyl
152	7	2.5	4688	2	F82885	hypothetical prote	225	6	2.2	129	2	I56195	gene Tap-1 protein
153	7	2.2	18	2	G42753	interferon alpha (226	6	2.2	130	2	PC4423	transcription regu
154	6	2.2	20	2	SS7202	vitronectin-bindin	227	6	2.2	133	2	T24099	hypothetical prote
155	6	2.2	21	2	A61413	interferon alpha (228	6	2.2	134	2	B83011	phosphoribosyl-AMP
156	6	2.2	22	2	B61413	interferon alpha (229	6	2.2	134	2	D39039	allergen Amb a 1 -
157	6	2.2	22	2	S09083	proteasome chain 2	230	6	2.2	134	2	D64381	conserved hypotet
158	6	2.2	33	2	I68894	gene TAP1 protein	231	6	2.2	135	2	D81243	hypothetical prote
159	6	2.2	33	2	D42753	interferon alpha (232	6	2.2	136	2	T68814	hypothetical prote
160	6	2.2	49	2	SS5324	endo-beta-1,6-gluc	233	6	2.2	137	2	A70909	hypothetical prote
161	6	2.2	50	2	H90982	H repeat-associate	234	6	2.2	142	2	T46542	hypothetical prote
162	6	2.2	61	1	B34913	adipokinetic hormo	235	6	2.2	143	2	T51530	40S RIBOSOMAL PROT
163	6	2.2	61	2	A27749	DNA-binding protei	236	6	2.2	143	2	S20937	photosystem I chai
164	6	2.2	62	2	A11756	hypothetical prote	237	6	2.2	143	2	A85333	hypothetical prote
165	6	2.2	64	2	C27749	DNA-binding protei	238	6	2.2	144	2	AF0663	conserved hypotet
166	6	2.2	67	2	T16372	hypothetical prote	239	6	2.2	144	2	C71149	hypothetical prote
167	6	2.2	68	2	C83767	small acid-soluble	240	6	2.2	145	2	S11239	Ig heavy chain V r
168	6	2.2	68	2	H75273	conserved hypotet	241	6	2.2	146	1	HANK2	hemoglobin II - ar
169	6	2.2	69	2	A41903	recombinase homolo	242	6	2.2	146	1	GGNKTB	globin I - ark she
170	6	2.2	70	2	AC0287	conserved hypotet	243	6	2.2	147	2	S36144	homoglobin I - ark
171	6	2.2	72	1	QQVZ9	hypothetical prote	244	6	2.2	147	2	F81266	30S ribosomal prot
172	6	2.2	72	2	S15168	gas-vesicle protei	245	6	2.2	147	2	AF1081	50S ribosomal prot
173	6	2.2	72	2	E41903	recombinase Bin3 -	246	6	2.2	148	2	D87687	conserved hypotet
174	6	2.2	73	2	F83266	hypothetical prote	247	6	2.2	148	2	B95131	conserved domain p
175	6	2.2	76	2	T51499	hypothetical prote	248	6	2.2	148	2		

249	6	2.2	148	2	T31040	conserved hypotet	322	6	2.2	185	2	S71512	hypothetical prote
250	6	2.2	149	2	F65169	o149 protein - Esc	323	6	2.2	186	2	G90541	gtp-binding protei
251	6	2.2	149	2	B91206	hypothetical prote	324	6	2.2	188	2	D70367	conserved hypotet
252	6	2.2	149	2	D86052	hypothetical prote	325	6	2.2	188	2	AZ2530	hypothetical prote
253	6	2.2	149	2	A13293	4-hydroxybenzoyl-C	326	6	2.2	189	1	IVHUF	interferon alpha-I
254	6	2.2	151	2	H70976	probable ribosomal	327	6	2.2	189	1	IVHUA5	interferon alpha-I
255	6	2.2	151	2	D86638	conserved hypotet	328	6	2.2	189	1	IVHUA3	interferon alpha-I
256	6	2.2	152	2	H95153	conserved hypotet	329	6	2.2	189	1	IVHUA9	interferon alpha-I
257	6	2.2	153	2	S60363	nucleoside-diphosp	330	6	2.2	189	1	IVHUA0	interferon alpha-7
258	6	2.2	153	2	S72982	ribosomal protein	331	6	2.2	189	1	IVHUI6	interferon alpha-I
259	6	2.2	153	2	A95928	probable amine oxi	332	6	2.2	189	1	IVHUA4	interferon alpha-4
260	6	2.2	154	2	P00465	beta C protein - P	333	6	2.2	189	1	IVHUI8	interferon alpha-I
261	6	2.2	155	2	B71223	hypothetical prote	334	6	2.2	189	2	IS2347	interferon alpha-M
262	6	2.2	155	2	C75191	transcriptlon regu	335	6	2.2	189	2	IS4464	interferon-alpha-F
263	6	2.2	156	1	A49342	acetyl-CoA carboxy	336	6	2.2	189	2	IS3102	interferon-alpha-J
264	6	2.2	156	2	A64990	diheme cytochrome	337	6	2.2	189	2	IS1970	interferon precurs
265	6	2.2	156	2	D91015	cytochrome c-type	338	6	2.2	190	2	A99308	probable phosphos
266	6	2.2	156	2	F85859	conserved hypotet	339	6	2.2	191	2	T25201	hypothetical prote
267	6	2.2	156	2	B90508	hypothetical prote	340	6	2.2	192	2	S43563	Ro1H10.5 protein -
268	6	2.2	156	4	S59288	transcriptlon elon	341	6	2.2	192	2	S43611	competence transcr
269	6	2.2	157	2	AC0998	hypothetical prote	342	6	2.2	192	2	G83096	conserved hypotet
270	6	2.2	157	2	T23781	hypothetical prote	343	6	2.2	193	2	F70775	hypothetical prote
271	6	2.2	158	2	D82782	hypothetical prote	344	6	2.2	194	2	B75613	conserved hypotet
272	6	2.2	158	2	F89123	protein K07C11.5 l	345	6	2.2	194	2	C75420	hypothetical prote
273	6	2.2	158	2	H83350	hypothetical prote	346	6	2.2	195	2	G75635	hypothetical prote
274	6	2.2	158	2	H85620	hypothetical prote	347	6	2.2	196	2	AR3096	phenylacrylic acid
275	6	2.2	159	2	AD0349	conserved hypotet	348	6	2.2	196	2	D8190	decarboxylase (imp
276	6	2.2	159	2	AE1241	B. subtilis Yq2C p	349	6	2.2	197	2	F75188	hypothetical prote
277	6	2.2	160	2	S65978	spore coat protein	350	6	2.2	198	1	Q0ECX3	putative methylase
278	6	2.2	161	1	F64447	conserved hypotet	351	6	2.2	198	2	B91168	hypothetical prote
279	6	2.2	162	2	C25843	interferon alpha-B	352	6	2.2	198	2	B86014	hypothetical prote
280	6	2.2	163	2	S61520	hemoglobin alpha-B	353	6	2.2	198	2	AF0991	conserved hypotet
281	6	2.2	165	2	S35287	eaf protein - phag	354	6	2.2	198	2	D72538	prphable [acyl-car
282	6	2.2	165	2	H64038	hypothetical prote	355	6	2.2	199	2	A69839	adenylylsulfate Ki
283	6	2.2	167	2	E25843	interferon alpha-F	356	6	2.2	199	2	I64125	nucleotide-binding
284	6	2.2	167	2	F25843	interferon alpha-J	357	6	2.2	199	2	D75442	probable peptidase
285	6	2.2	167	2	S49628	FUN81 protein - ye	358	6	2.2	200	1	RDMD4	ribosomal protein
286	6	2.2	168	2	AG3296	transcriptlon regu	359	6	2.2	200	2	B95349	probable decarboxy
287	6	2.2	168	2	B85740	hypothetical prote	360	6	2.2	200	2	F64068	recombination prot
288	6	2.2	168	2	E90878	probable filament	361	6	2.2	201	2	G69340	cobalamin biosynth
289	6	2.2	168	2	T17204	hypothetical prote	362	6	2.2	202	2	S68609	recombinase Sin -
290	6	2.2	168	2	G95023	acetyltransferase,	363	6	2.2	202	2	D59662	probable metal-dep
291	6	2.2	168	2	H97894	hypothetical prote	364	6	2.2	203	2	F90532	conserved hypotet
292	6	2.2	168	2	E84362	GTP cyclohydrolase	365	6	2.2	203	2	T19760	hypothetical prote
293	6	2.2	170	2	A65136	transcriptlon elon	366	6	2.2	204	2	G40661	riva protein - Hae
294	6	2.2	170	2	H91159	transcriptlon elon	367	6	2.2	204	2	AE3302	multiple antibioti
295	6	2.2	170	2	G86005	transcriptlon elon	368	6	2.2	205	2	E70199	competence protein
296	6	2.2	170	2	B95169	conserved domain p	369	6	2.2	206	2	AC3478	protein ybis precu
297	6	2.2	170	2	D75400	conserved hypotet	370	6	2.2	207	2	A10005	guanylate kinase (
298	6	2.2	170	2	C98035	hypothetical prote	371	6	2.2	207	2	T31959	hypothetical prote
299	6	2.2	171	2	S77242	hypothetical prote	372	6	2.2	207	2	T51567	hypothetical prote
300	6	2.2	173	2	T24942	hypothetical prote	373	6	2.2	208	1	R5Y4C	ribosomal protein
301	6	2.2	173	2	C86503	acetyltransferase (l	374	6	2.2	208	2	A38202	GTP-binding protei
302	6	2.2	176	2	I56314	interferon-alpha-	375	6	2.2	208	2	T33385	60S ribosomal prot
303	6	2.2	176	2	A72237	hypothetical prote	376	6	2.2	208	2	A38594	tropoin I - fruit
304	6	2.2	176	2	A83050	outer membrane lip	377	6	2.2	208	2	A40547	tropoin I - fruit
305	6	2.2	178	2	D71668	gripe protein (grip	378	6	2.2	209	2	F83149	thiamin-phosphate
306	6	2.2	178	2	A71939	hypothetical prote	379	6	2.2	209	2	D88577	protein Ro1H10.5 (
307	6	2.2	178	2	S64648	hypothetical prote	380	6	2.2	210	1	D64238	hypothetical prote
308	6	2.2	178	2	S67379	hypothetical prote	381	6	2.2	210	2	F71308	probable 2-dehydro
309	6	2.2	179	2	T49757	hypothetical prote	382	6	2.2	210	2	S76316	hypothetical prote
310	6	2.2	179	2	T51570	hypothetical prote	383	6	2.2	210	2	F83393	probable transcrip
311	6	2.2	180	2	T41322	probable succinate	384	6	2.2	210	2	E75349	hypothetical prote
312	6	2.2	180	2	E64834	probable type 1 fi	385	6	2.2	212	2	T36864	probable cyclohexa
313	6	2.2	180	2	S58767	streptothricin ace	386	6	2.2	212	2	C55544	flbE protein - Cau
314	6	2.2	180	2	H84154	hypothetical prote	387	6	2.2	213	2	B89820	hypothetical prote
315	6	2.2	180	2	D90659	hypothetical prote	388	6	2.2	213	2	P98019	hypothetical prote
316	6	2.2	181	2	I56313	interferon alpha 2	389	6	2.2	214	2	B95898	probable sensory t
317	6	2.2	181	2	E82077	ampd protein VC242	390	6	2.2	215	2	S61337	proteasome endopep
318	6	2.2	181	2	T35851	hypothetical prote	391	6	2.2	216	2	S27382	hypothetical prote
319	6	2.2	182	2	G70698	peptidylprolyl iso	392	6	2.2	217	2	B97948	hypothetical prote
320	6	2.2	185	2	S67036	SPF protein - yea	393	6	2.2	218	2	H64437	trk system potassi
321	6	2.2	185	2	C82662	hypothetical prote	394	6	2.2	218	2	B64365	hypothetical prote

395	6	2.2	219	2	D82603	hypotheical prote
396	6	2.2	220	2	H69257	hypotheical prote
397	6	2.2	220	2	T35744	hypotheical prote
398	6	2.2	221	2	A39491	conserved hypothe
399	6	2.2	222	2	A56276	carbon monoxide ox
400	6	2.2	225	2	T11471	H+-transporting tw
401	6	2.2	226	2	S10496	hypotheical prote
402	6	2.2	226	2	AG1883	hypotheical prote
403	6	2.2	227	2	AB1076	probable fibrillar
404	6	2.2	228	2	T47999	hypotheical prote
405	6	2.2	228	2	E69517	hypotheical prote
406	6	2.2	228	2	F70934	probable lpgn prot
407	6	2.2	229	2	E71010	hypotheical prote
408	6	2.2	229	2	T15359	hypotheical prote
409	6	2.2	230	2	A61607	probable hemolysin
410	6	2.2	231	2	G72463	hypotheical prote
411	6	2.2	231	2	J01601	replication protei
412	6	2.2	231	2	C70000	ABC transporter (A
413	6	2.2	232	1	DXCH	ovalbumin-related
414	6	2.2	232	2	A95081	hypotheical prote
415	6	2.2	237	2	F82178	probable fibrillar
416	6	2.2	237	2	C97948	hypotheical prote
417	6	2.2	239	2	A12773	sugar fermentation
418	6	2.2	239	2	G97553	hypotheical prote
419	6	2.2	239	2	TS1565	hypotheical prote
420	6	2.2	239	2	G90366	hypotheical prote
421	6	2.2	239	2	AH2836	O-methyltransfera
422	6	2.2	239	2	B97614	hypotheical prote
423	6	2.2	240	2	G91045	hypotheical prote
424	6	2.2	240	2	C85890	hypotheical prote
425	6	2.2	240	2	H65002	hypotheical prote
426	6	2.2	241	2	AP0758	precorrin-3B C17-m
427	6	2.2	241	2	S19267	anthranilate phosp
428	6	2.2	242	2	T37107	hypotheical prote
429	6	2.2	242	2	AB1091	autolysin, N-acety
430	6	2.2	242	2	A11454	arginine-binding p
431	6	2.2	243	2	AG0607	probable RNA methy
432	6	2.2	243	2	AH0824	capsula polysacch
433	6	2.2	243	2	UC5723	conserved membrane
434	6	2.2	243	2	C96909	hypotheical prote
435	6	2.2	244	2	T20784	hypotheical prote
436	6	2.2	244	2	A90302	oxidizing 5'-phosp
437	6	2.2	245	2	AP0271	phosphatase - cat
438	6	2.2	245	2	S52096	probable ATP synth
439	6	2.2	245	2	F91053	probable ATP synth
440	6	2.2	245	2	B85888	conserved hypothe
441	6	2.2	245	2	S66095	hypotheical prote
442	6	2.2	245	2	C83584	major antigenic pe
443	6	2.2	245	2	B81349	dipeptide transpor
444	6	2.2	245	2	A99163	hypotheical prote
445	6	2.2	245	2	AH3124	hypotheical prote
446	6	2.2	246	1	C65030	probable membrane
447	6	2.2	246	2	B81305	hypotheical prote
448	6	2.2	246	2	H87427	hypotheical prote
449	6	2.2	246	2	F97076	activin type I rec
450	6	2.2	247	2	PC4260	hypotheical prote
451	6	2.2	247	2	T26461	conserved hypothe
452	6	2.2	248	2	B95203	conserved hypothe
453	6	2.2	248	2	B98070	60S ribosomal prot
454	6	2.2	249	1	R5B77	surfeit locus prot
455	6	2.2	249	2	AF3434	uncharacterized pt
456	6	2.2	249	2	A97154	hypotheical prote
457	6	2.2	249	2	T21920	probable UDP-N-ac
458	6	2.2	249	2	F71342	superoxide dismuta
459	6	2.2	251	2	S52989	lactose operon rep
460	6	2.2	251	2	A44506	conserved hypothe
461	6	2.2	251	2	G90015	hypotheical prote
462	6	2.2	251	2	F83101	hypotheical prote
463	6	2.2	251	2	A84367	hypotheical prote
464	6	2.2	252	2	C89825	dehydrogenase homo
465	6	2.2	252	2	AG1346	ABC transporter, A
466	6	2.2	253	2	E69497	gpi2 protein - Myc
467	6	2.2	254	2	A72801	
468	6	2.2	254	2	B81293	hypotheical prote
469	6	2.2	255	2	B81929	probable imidazole
470	6	2.2	255	2	H81176	hlf protein NMB06
471	6	2.2	255	2	S74930	hypotheical prote
472	6	2.2	255	2	DC4833	probable ABC-type
473	6	2.2	255	2	F85619	probable ABC-type
474	6	2.2	255	2	H90755	probable ABC-type
475	6	2.2	256	2	G75404	outer membrane pro
476	6	2.2	256	2	H75404	outer membrane pro
477	6	2.2	256	2	T35795	hypotheical prote
478	6	2.2	257	2	AB0353	probable Spou-fam1
479	6	2.2	258	2	I16945	haptoglobin Hsp - C
480	6	2.2	258	2	I16947	haptoglobin Hsp - C
481	6	2.2	258	2	G02959	haptoglobin - fms
482	6	2.2	258	2	F84812	similar to Mtn3 pr
483	6	2.2	259	1	B64251	probable methyltra
484	6	2.2	259	2	PN0687	cyclin - anthracno
485	6	2.2	259	2	G83153	conserved hypothe
486	6	2.2	259	2	CC4437	hypotheical prote
487	6	2.2	259	2	AB3238	hypotheical prote
488	6	2.2	259	2	AG8302	oligopeptide trans
489	6	2.2	259	2	AG2981	hypotheical prote
490	6	2.2	260	1	A27058	chlorocatechol 1,2
491	6	2.2	260	2	T44616	chlorocatechol 1,2
492	6	2.2	260	2	T44666	chlorocatechol 1,2
493	6	2.2	260	2	B90026	hypotheical prote
494	6	2.2	261	2	H83442	probable enoyl-CoA
495	6	2.2	262	2	B42478	glutamine-binding
496	6	2.2	262	2	B81326	probable periplasm
497	6	2.2	262	2	H87217	probable undecapre
498	6	2.2	262	2	T14807	hypotheical prote
499	6	2.2	264	2	F83351	conserved hypothe
500	6	2.2	264	2	H69807	hypotheical prote
501	6	2.2	264	2	AB1362	probable terminase
502	6	2.2	265	2	G84049	hypotheical prote
503	6	2.2	265	2	C64579	molycopdenum ABC tra
504	6	2.2	265	2	G71934	molycopdenum ABC tra
505	6	2.2	267	2	T24397	hypotheical prote
506	6	2.2	267	2	A69867	conserved hypothe
507	6	2.2	267	2	A75217	hypotheical prote
508	6	2.2	267	2	D95089	hypotheical prote
509	6	2.2	267	2	G97956	peptidyl-prolyl ci
510	6	2.2	268	2	A71952	dipeptide transpor
511	6	2.2	268	2	F64557	dipeptide transpor
512	6	2.2	268	2	C87579	romB-dependent rec
513	6	2.2	268	2	AD3496	ABC transporter su
514	6	2.2	268	2	S45091	hypotheical prote
515	6	2.2	268	2	AB2458	hypotheical prote
516	6	2.2	269	1	H75003	beta-lactamase reg
517	6	2.2	269	2	H75576	cobalam synthase
518	6	2.2	271	2	AD0441	probable aliphatic
519	6	2.2	272	1	IPEC28	lipoprotein-28 pre
520	6	2.2	272	2	E70879	probable dehydroge
521	6	2.2	272	2	G91203	lipoprotein-28 (im
522	6	2.2	272	2	DB6049	lipoprotein-28 (im
523	6	2.2	272	2	G97021	lipoprotein, attac
524	6	2.2	273	2	T51977	proteasome endope
525	6	2.2	273	2	H86370	3.2 kD hypotheical
526	6	2.2	274	2	T51979	proteasome endope
527	6	2.2	274	2	GB1914	hypotheical prote
528	6	2.2	274	2	H81187	conserved hypothe
529	6	2.2	274	2	GB3214	probable ATP-bind
530	6	2.2	274	2	C66213	hypotheical prote
531	6	2.2	274	2	A82509	probable phosphati
532	6	2.2	275	2	T42747	hypotheical prote
533	6	2.2	275	2	S76778	hypotheical prote
534	6	2.2	275	2	DB3626	transcription regu
535	6	2.2	275	2	B83983	hypotheical prote
536	6	2.2	276	2	H83368	probable permease
537	6	2.2	276	2	H75169	abc transporter PA
538	6	2.2	276	2	AG3305	23S ribosomal RNA
539	6	2.2	277	2	C93358	probable oxidoredu
540	6	2.2	277	2	G71849	amino acid ABC tra

541	6	2.2	277	2	D64666	614	6	2.2	308	2	AF1581	B. subtilis ribonu
542	6	2.2	278	2	B82388	615	6	2.2	309	1	XYECM	homoserine O-succi
543	6	2.2	279	2	S27149	616	6	2.2	309	1	XYECM	carbamate kinase (
544	6	2.2	279	2	A87108	617	6	2.2	309	2	C01245	homoserine transu
545	6	2.2	280	2	PH0269	618	6	2.2	309	2	A86093	homoserine transu
546	6	2.2	280	2	A89612	619	6	2.2	309	2	AC1011	homoserine O-succi
547	6	2.2	280	2	A83796	620	6	2.2	309	2	AD0502	homoserine kinase
548	6	2.2	280	2	D75077	621	6	2.2	309	2	A12809	hypothetical prote
549	6	2.2	282	2	H84653	622	6	2.2	309	2	A90439	ABC transporter, A
550	6	2.2	283	2	JC5661	623	6	2.2	309	2	T52453	ATP-dependent Clp
551	6	2.2	284	2	B95017	624	6	2.2	310	2	CE9218	transcription init
552	6	2.2	284	2	C97890	625	6	2.2	310	2	T52041	probable ATP-depen
553	6	2.2	285	1	C70873	626	6	2.2	311	2	D64209	hydroxymethylgluta
554	6	2.2	285	2	A95851	627	6	2.2	312	2	AC2904	oxidoreductase Atu
555	6	2.2	286	2	JC6007	628	6	2.2	312	2	E97679	probable oxidoredu
556	6	2.2	286	2	AH1596	629	6	2.2	313	2	AD0920	porphobilinogen de
557	6	2.2	286	2	AH1596	630	6	2.2	313	2	CE9768	hypothetical prote
558	6	2.2	286	2	G84898	631	6	2.2	313	2	A59496	trp domain-contain
559	6	2.2	286	2	E70124	632	6	2.2	314	1	DEGXM	malate dehydrogena
560	6	2.2	287	2	C81023	633	6	2.2	314	1	A41671	iron transport pro
561	6	2.2	287	2	G81968	634	6	2.2	315	1	YXCKTA	thymidylate syntha
562	6	2.2	287	2	AD0006	635	6	2.2	315	2	H96598	protein F20N2.8 [i
563	6	2.2	289	2	S12619	636	6	2.2	316	2	JC4701	cadmium, zinc, cob
564	6	2.2	289	2	AB1862	637	6	2.2	316	2	D71375	probable ABC trans
565	6	2.2	291	2	E84408	638	6	2.2	316	2	A41626	protein farnesyltr
566	6	2.2	291	2	F69959	639	6	2.2	317	2	E95239	conserved hypotnet
567	6	2.2	292	2	T16085	640	6	2.2	317	2	E98103	conserved hypotnet
568	6	2.2	293	2	H65182	641	6	2.2	317	2	E86419	probable exonuclea
569	6	2.2	293	2	B91219	642	6	2.2	318	2	T7548	probable lacyl-car
570	6	2.2	293	2	D86065	643	6	2.2	318	2	G82086	homoserine kinase
571	6	2.2	293	2	D81204	644	6	2.2	318	2	D96774	unknown protein Fl
572	6	2.2	293	2	A81780	645	6	2.2	318	2	T00402	homocodomain transc
573	6	2.2	293	2	D95919	646	6	2.2	320	2	G95268	probable ABC trans
574	6	2.2	293	2	T31146	647	6	2.2	321	2	T38413	probable oxidoredu
575	6	2.2	294	1	G65106	648	6	2.2	321	2	G96921	oligopeptide ABC t
576	6	2.2	294	2	G72293	649	6	2.2	322	2	S74505	erthrococyte band 7
577	6	2.2	294	2	T10690	650	6	2.2	322	2	A83857	3-dehydroquinare s
578	6	2.2	294	2	D91134	651	6	2.2	322	2	S52032	triose-phosphate i
579	6	2.2	294	2	AB0424	652	6	2.2	322	2	G64388	hypothetical prote
580	6	2.2	294	2	G85979	653	6	2.2	322	2	G72643	hypothetical prote
581	6	2.2	294	2	AC0901	654	6	2.2	323	2	A64054	selenophosphate sy
582	6	2.2	296	1	JQ2155	655	6	2.2	323	2	D83708	phosphonates trans
583	6	2.2	296	1	T51326	656	6	2.2	324	2	AC0088	putative flagellar
584	6	2.2	296	2	T46619	657	6	2.2	325	2	F82068	2-hydroxyacid dehy
585	6	2.2	297	2	T22690	658	6	2.2	325	2	T07001	cysteine synthase
586	6	2.2	297	2	T34165	659	6	2.2	325	2	C72289	oligopeptide ABC t
587	6	2.2	298	1	C42802	660	6	2.2	326	2	T25728	hypothetical prote
588	6	2.2	298	1	D69023	661	6	2.2	326	2	B69412	transcription init
589	6	2.2	298	2	B83553	662	6	2.2	327	2	B67168	probable membrane
590	6	2.2	298	2	T24029	663	6	2.2	328	2	I64164	hypothetical prote
591	6	2.2	298	2	D95987	664	6	2.2	329	1	HPDG	haptoglobin precur
592	6	2.2	299	2	A95878	665	6	2.2	329	2	AC0942	periplasmic sulpha
593	6	2.2	299	2	T37556	666	6	2.2	330	2	S67139	probable membrane
594	6	2.2	300	2	B64308	667	6	2.2	330	2	C83995	branched-chain alp
595	6	2.2	300	2	H87390	668	6	2.2	330	2	C69593	3-methyl-2-oxobuta
596	6	2.2	300	2	F69719	669	6	2.2	332	2	C97111	glycerol 3-phospha
597	6	2.2	300	2	AC3175	670	6	2.2	332	2	A46408	abscisic acid-indu
598	6	2.2	301	2	AB3195	671	6	2.2	332	2	A36653	involued in fatty
599	6	2.2	301	2	S51132	672	6	2.2	333	2	H69679	oligopeptide trans
600	6	2.2	302	2	A89898	673	6	2.2	333	2	AE0266	hemagglutinin - me
601	6	2.2	302	2	A40369	674	6	2.2	333	2	P00377	oligopeptide trans
602	6	2.2	303	1	GEPCO	675	6	2.2	334	2	OREBOF	probable aspartate
603	6	2.2	303	2	H71693	676	6	2.2	334	2	H71524	oligopeptide trans
604	6	2.2	304	2	E97343	677	6	2.2	334	2	B64872	hypothetical prote
605	6	2.2	305	2	B97762	678	6	2.2	334	2	B85705	hypothetical prote
606	6	2.2	305	2	A770482	679	6	2.2	334	2	C90847	oligopeptide trans
607	6	2.2	305	2	F86456	680	6	2.2	334	2	AC0651	oligopeptide trans
608	6	2.2	306	2	E83169	681	6	2.2	335	2	F71211	probable oligopept
609	6	2.2	306	2	T46757	682	6	2.2	335	2	F75026	oligopeptide trans
610	6	2.2	306	2	AG2104	683	6	2.2	335	2	AG0183	galactose-binding
611	6	2.2	306	2	S27496	684	6	2.2	335	2	G81280	holiday junction
612	6	2.2	308	2	S24202	685	6	2.2	335	2	A41845	off A - Treponema
613	6	2.2	308	2	AD1228	686	6	2.2	335	2	A72289	oligopeptide ABC t

687	6	2.2	335	2	AD3492	760	6	2.2	354	2	T52401	branched-chain ami
688	6	2.2	336	1	A64500	761	6	2.2	354	2	C90882	probable transcript
689	6	2.2	336	2	F82242	762	6	2.2	354	2	F85736	probable transcript
690	6	2.2	336	2	F71942	763	6	2.2	354	2	A64894	probable regulator
691	6	2.2	336	2	C64652	764	6	2.2	354	2	A12889	lysyl-tRNA synthet
692	6	2.2	336	2	E84536	765	6	2.2	354	2	F97665	NodI membrane tran
693	6	2.2	336	2	T00832	766	6	2.2	355	2	H85320	heat inducible tra
694	6	2.2	336	2	T21565	767	6	2.2	355	2	E97928	NADH dehydrogenase
695	6	2.2	336	2	H69437	768	6	2.2	356	2	D87450	coenzyme PQQ synth
696	6	2.2	337	2	C70990	769	6	2.2	356	2	D84274	probable sulfite o
697	6	2.2	338	1	DEPTMM	770	6	2.2	357	2	B75575	probable aspartate
698	6	2.2	338	1	DEMSM	771	6	2.2	357	2	T29856	hypothetical prote
699	6	2.2	338	2	T03432	772	6	2.2	357	2	S76088	probable triacylgly
700	6	2.2	338	2	F83965	773	6	2.2	357	2	T01607	hypothetical prote
701	6	2.2	338	2	A83835	774	6	2.2	357	2	T27919	protein co-factor
702	6	2.2	339	2	A41677	775	6	2.2	358	2	I46532	decorin precursor
703	6	2.2	339	2	AC3245	776	6	2.2	359	1	NEHUC8	hypothetical prote
704	6	2.2	339	2	T20442	777	6	2.2	359	2	T21705	conserved hypochet
705	6	2.2	339	2	T27715	778	6	2.2	359	2	AD2743	hypothetical prote
706	6	2.2	340	2	F70616	779	6	2.2	359	2	C97524	decorin precursor
707	6	2.2	340	2	JH0363	780	6	2.2	360	2	S06280	decorin
708	6	2.2	340	2	B41969	781	6	2.2	360	2	I47020	hypothetical prote
709	6	2.2	340	2	E69544	782	6	2.2	360	2	H84401	hypothetical prote
710	6	2.2	340	2	D98346	783	6	2.2	361	2	T39784	hypothetical prote
711	6	2.2	341	2	G71313	784	6	2.2	361	2	AC2936	glycerol dehydroge
712	6	2.2	341	2	F87620	785	6	2.2	362	2	G95029	glycerol dehydroge
713	6	2.2	342	2	A30189	786	6	2.2	362	2	B97901	phosphoserine amin
714	6	2.2	342	2	S75086	787	6	2.2	362	2	C82572	probable murein pep
715	6	2.2	342	2	S57510	788	6	2.2	363	2	G95937	hypothetical prote
716	6	2.2	342	2	C71372	789	6	2.2	364	2	G84367	conserved hypochet
717	6	2.2	342	2	A83667	790	6	2.2	365	2	D89960	prephenate dehydro
718	6	2.2	343	2	D91032	791	6	2.2	367	2	T17481	conserved hypochet
719	6	2.2	343	2	E85976	792	6	2.2	367	2	G83309	VSG expression sit
720	6	2.2	344	2	S42648	793	6	2.2	368	2	JH0348	hypothetical prote
721	6	2.2	344	2	AB2306	794	6	2.2	368	2	G86416	dihydroxotriate oxi
722	6	2.2	344	2	S76268	795	6	2.2	369	1	A23559	larnesyltransferase
723	6	2.2	344	2	T05524	796	6	2.2	369	1	S53722	hypothetical prote
724	6	2.2	344	2	H95059	797	6	2.2	369	2	S76784	hypothetical prote
725	6	2.2	344	2	AG1005	798	6	2.2	369	2	D87476	hypothetical prote
726	6	2.2	345	2	F95216	799	6	2.2	369	2	S75038	hypothetical prote
727	6	2.2	345	2	D98080	800	6	2.2	369	2	T16506	glycerol dehydroge
728	6	2.2	345	2	I36941	801	6	2.2	370	1	U01474	(R,R)-butanediol d
729	6	2.2	345	2	S09858	802	6	2.2	370	2	A86737	glibberellin 20-oxi
730	6	2.2	345	2	A10830	803	6	2.2	370	2	T11849	8-amino-7-oxononan
731	6	2.2	345	2	C95873	804	6	2.2	370	2	B72018	oxononanoate synth
732	6	2.2	345	2	AH3131	805	6	2.2	370	2	A86506	branched chain ami
733	6	2.2	345	2	AD3635	806	6	2.2	370	2	C72290	probable N-acetylm
734	6	2.2	345	2	C90416	807	6	2.2	371	2	D75402	galactose-1-phosph
735	6	2.2	345	2	G83633	808	6	2.2	372	2	G82181	hypothetical prote
736	6	2.2	345	2	T19834	809	6	2.2	372	2	AB2061	sarcosine oxidase
737	6	2.2	346	2	I36942	810	6	2.2	373	2	H84187	hypothetical prote
738	6	2.2	346	2	C81446	811	6	2.2	373	2	B84595	hypothetical prote
739	6	2.2	346	2	C98156	812	6	2.2	373	2	A69773	hypothetical prote
740	6	2.2	346	2	T19676	813	6	2.2	375	2	T39364	probable galactosyl
741	6	2.2	347	1	HPHUT	814	6	2.2	375	2	C95106	carboxymorspermid
742	6	2.2	347	1	HPMS	815	6	2.2	375	2	A13383	hypothetical prote
743	6	2.2	347	1	HPRT	816	6	2.2	375	2	T32251	hypothetical prote
744	6	2.2	347	1	F90937	817	6	2.2	375	2	E87974	carboxymorspermid
745	6	2.2	347	2	B85786	818	6	2.2	376	2	T02770	glutathione-disulf
746	6	2.2	347	2	JM0033	819	6	2.2	376	2	E35879	probable sugar ABC
747	6	2.2	347	2	AH0710	820	6	2.2	377	2	F97250	proline/glycine be
748	6	2.2	347	2	T38154	821	6	2.2	377	2	A38662	tRNA-splicing endo
749	6	2.2	347	2	B96008	822	6	2.2	377	2	D66881	pyridine nucleotid
750	6	2.2	347	2	D97588	823	6	2.2	378	1	OKGARI	protein kinase (EC
751	6	2.2	348	2	HPHUR	824	6	2.2	378	2	C70119	hypothetical prote
752	6	2.2	349	2	I36944	825	6	2.2	380	2	T00683	hypothetical prote
753	6	2.2	350	2	A42009	826	6	2.2	380	2	T48052	hypothetical prote
754	6	2.2	350	2	F75448	827	6	2.2	381	2	S74777	hypothetical prote
755	6	2.2	352	1	CHECFX	828	6	2.2	381	2	JC4378	metalloproteinase
756	6	2.2	352	1	G91167	829	6	2.2	381	2	T38655	hypothetical 37.5
757	6	2.2	352	1	G86013	830	6	2.2	383	2	S11433	ovalbumin - Japane
758	6	2.2	353	2	T03687	831	6	2.2	385	2	G87340	antioctaniferase
759	6	2.2	353	2	A10230	832	6	2.2	385	2	A44102	di-N-acetylchitobi

833	6	2.2	385	1	OACH	ovalbumin (validat	906	6	2.2	417	2	A13382	lysine-tRNA ligase
834	6	2.2	387	2	C92336	conserved hypochet	907	6	2.2	417	2	S74940	hypothetical prote
835	6	2.2	387	2	A83072	RND multidrug effl	908	6	2.2	417	2	S56454	hypothetical ABC t
836	6	2.2	387	2	T28402	ORF MSV41 leucine	909	6	2.2	418	2	C83341	serine hydroxymeth
837	6	2.2	388	1	DYCH	ovalbumin-related	910	6	2.2	419	2	AF0104	maltopectin [import
838	6	2.2	388	2	B48899	beta-lactamase (EC	911	6	2.2	419	2	E95157	Atz/712 family pro
839	6	2.2	388	2	T26977	hypothetical prote	912	6	2.2	420	1	AUB1RS	argininosuccinate
840	6	2.2	388	2	AC0138	tolA colicin impor	913	6	2.2	420	2	T34750	serine hydroxymeth
841	6	2.2	389	2	TS0671	sarcosine oxidase	914	6	2.2	420	2	A96535	unknown protein, 1
842	6	2.2	390	2	AF0426	glucose-6-phosphat	915	6	2.2	420	2	E96614	hypothetical prote
843	6	2.2	390	2	AC0112	hypothetical prote	916	6	2.2	421	1	D64100	glycine hydroxymet
844	6	2.2	392	2	D53240	allergen Amb a 1.4	917	6	2.2	421	2	JY0057	tolA protein -Bac
845	6	2.2	392	2	D86823	cell division prot	918	6	2.2	422	2	G84073	C4-dicarboxylate t
846	6	2.2	393	2	C83049	probable two-compo	919	6	2.2	423	2	S86808	hypothetical prote
847	6	2.2	393	2	A71154	hypothetical prote	920	6	2.2	424	2	F95894	probable nitrate t
848	6	2.2	393	2	D97275	glycosyltransferas	921	6	2.2	424	2	A95159	psr protein [import
849	6	2.2	393	2	G84094	phosphoglycerate k	922	6	2.2	425	2	A98025	conserved hypochet
850	6	2.2	394	2	F90725	membrane spanning	923	6	2.2	425	2	S53004	mitosis-specific c
851	6	2.2	394	2	G85576	alcohol dehydrogen	924	6	2.2	426	2	F71871	topoisomerase I -
852	6	2.2	395	2	E72381	probable cysteine	925	6	2.2	426	2	C70896	glycine hydroxymet
853	6	2.2	395	2	T38685	probable amidohydr	926	6	2.2	428	2	B87463	hypothetical prote
854	6	2.2	395	2	D81408	allergen Amb a 1.1	927	6	2.2	428	2	D70583	probable transmemb
855	6	2.2	396	2	A39099	glucose-6-phosphat	928	6	2.2	428	2	B89902	hypothetical prote
856	6	2.2	396	2	T04100	hypothetical prote	929	6	2.2	429	2	T45520	transposase [import
857	6	2.2	396	2	A84556	hypothetical prote	930	6	2.2	429	2	S29044	endothiase A pr
858	6	2.2	396	2	AE2137	allergen Amb a 1.3	931	6	2.2	430	2	AD2873	hypothetical prote
859	6	2.2	397	2	C53240	allergen Amb a 1.3	932	6	2.2	430	2	G70352	N-ethylameline ch
860	6	2.2	397	2	C39099	conserved hypochet	933	6	2.2	430	2	S55325	endo-beta-1,6-gluc
861	6	2.2	397	2	A64014	hypothetical prote	934	6	2.2	432	2	D55333	hypothetical prote
862	6	2.2	397	2	T19182	aminotransferase p	935	6	2.2	433	2	E96649	hypothetical prote
863	6	2.2	397	2	A84064	phosphoglycerate k	936	6	2.2	435	2	A81416	probable integral
864	6	2.2	398	2	E86655	serin - barley	937	6	2.2	435	2	E64079	putrescine/cornubi
865	6	2.2	398	2	S29819	hypothetical prote	938	6	2.2	436	2	T01805	hypothetical prote
866	6	2.2	398	2	T34847	allergen Amb a 1.2	939	6	2.2	436	2	E64340	hypothetical prote
867	6	2.2	398	2	B53340	allergen Amb a 1.2	940	6	2.2	437	2	T39883	hypothetical prote
868	6	2.2	398	2	B39099	glycosyltransferas	941	6	2.2	437	1	D64457	translational initia
869	6	2.2	398	2	C71728	probable oxygen-in	942	6	2.2	437	2	I39578	nccC protein - Alc
870	6	2.2	399	2	T28359	ORF MSV198 Mtg mot	943	6	2.2	437	2	H96933	hypothetical prote
871	6	2.2	399	2	A47094	aspartate transam	944	6	2.2	439	2	D70958	hypothetical prote
872	6	2.2	400	2	D86575	phosphoglycerate k	945	6	2.2	439	2	B86302	hypothetical prote
873	6	2.2	402	2	T72949	phosphate/triose-p	946	6	2.2	440	2	H70943	hypothetical prote
874	6	2.2	402	2	T14438	hypothetical prote	947	6	2.2	440	2	IS5442	peroxisome prolif
875	6	2.2	402	2	B98175	hypothetical prote	948	6	2.2	440	2	UC4530	probable MFS trans
876	6	2.2	403	2	C70832	hypothetical prote	949	6	2.2	441	2	C83368	probable DNA-damag
877	6	2.2	403	2	H96026	probable amino aci	950	6	2.2	441	2	AE1015	hypothetical prote
878	6	2.2	405	2	S64052	26S proteasome reg	951	6	2.2	443	1	A46248	methy1 coenzyme M
879	6	2.2	405	2	C83204	hypothetical prote	952	6	2.2	443	2	S43899	N-ethylameline ch
880	6	2.2	405	2	E81946	hypothetical prote	953	6	2.2	445	2	S83743	hypothetical prote
881	6	2.2	405	2	F82229	hapoglobin precur	954	6	2.2	445	2	T28015	muslin A, tracheal
882	6	2.2	406	1	HPH02	argininosuccinate	955	6	2.2	445	2	E66440	hypothetical prote
883	6	2.2	406	2	C81415	hypothetical prote	956	6	2.2	445	2	F97649	hypothetical prote
884	6	2.2	407	2	H81160	hypothetical prote	957	6	2.2	446	2	T35005	probable integral
885	6	2.2	408	2	C91062	hypothetical prote	958	6	2.2	447	2	S32227	glutamate dehydrog
886	6	2.2	408	2	G85905	hypothetical prote	959	6	2.2	447	2	T02547	hypothetical prote
887	6	2.2	408	2	G65038	hypothetical prote	960	6	2.2	447	2	G96569	hypothetical prote
888	6	2.2	408	2	F69614	fosmidmycin resist	961	6	2.2	447	2	D85646	hypothetical prote
889	6	2.2	409	1	C95971	probable sugar upt	962	6	2.2	447	2	F90786	hypothetical prote
890	6	2.2	410	2	H95888	hypothetical prote	963	6	2.2	447	2	T18264	cellulose anhor
891	6	2.2	410	2	AH1484	probable cell surf	964	6	2.2	448	2	G70009	organic acid trans
892	6	2.2	410	2	AE3092	hypothetical prote	965	6	2.2	448	2	H97249	protein containing
893	6	2.2	410	2	E98194	hypothetical prote	966	6	2.2	449	2	AF1339	phosphoglucomutase
894	6	2.2	410	2	G30315	polyterredoxin 6x2	967	6	2.2	450	2	H82044	C4-dicarboxylate t
895	6	2.2	412	1	E70108	aminoputridase II	968	6	2.2	451	2	B96495	hypothetical prote
896	6	2.2	412	2	S62538	hypothetical coll	969	6	2.2	451	2	S57588	phosphomevalonate
897	6	2.2	412	2	T41552	hypothetical prote	970	6	2.2	451	2	T30732	A-type inclusion b
898	6	2.2	412	2	S66925	hypothetical prote	971	6	2.2	451	2	B96002	hypothetical prote
899	6	2.2	414	2	AI0156	probable cryptopha	972	6	2.2	454	2	B86121	conserved hypochet
900	6	2.2	414	2	T34815	two-component syst	973	6	2.2	454	2	S56459	probable ligase [i
901	6	2.2	414	2	AG0075	probable sugar bin	974	6	2.2	457	2	B86121	UDP-N-acetylmutama
902	6	2.2	416	2	AG2892	hypothetical prote	975	6	2.2	457	2	S56459	muirein peptide lig
903	6	2.2	416	2	G82968	serine hydroxymeth	976	6	2.2	457	2	AC1056	
904	6	2.2	417	2			977	6	2.2	457	2		
905	6	2.2	417	2			978	6	2.2	457	2		

979	6	2.2	457	2	T25612	hypothetical prote
980	6	2.2	457	2	S75609	hypothetical prote
981	6	2.2	457	2	T08861	hypothetical prote
982	6	2.2	459	2	A13384	biotin carboxylase
983	6	2.2	459	2	D10731	hypothetical prote
984	6	2.2	461	2	B81862	probable integral
985	6	2.2	461	2	B81080	drug resistance tr
986	6	2.2	461	2	B83134	probable aminotran
987	6	2.2	461	2	A97789	hypothetical prote
988	6	2.2	462	2	A97782	phosphomannomutase
989	6	2.2	462	2	B81551	lipid A biosynthes
990	6	2.2	462	2	B84689	chloroplast membra
991	6	2.2	462	2	B83123	probable endoprote
992	6	2.2	462	2	C83660	cell-cycle protein
993	6	2.2	463	2	D89337	glycyl-tRNA synthe
994	6	2.2	463	2	A64381	NADH oxidase - Met
995	6	2.2	464	2	G97290	cysteinyI-CRNAsyn
996	6	2.2	464	2	H90340	hypothetical prote
997	6	2.2	466	2	S36209	dep protein precur
998	6	2.2	466	2	T28396	ORF MSV235 probabl
999	6	2.2	467	2	AD3270	dihydrolipoamide d
1000	6	2.2	467	2	B72119	acyltransferase -

ALIGNMENTS

RESULT 1

Outer membrane 30K protein - Pasteurella haemolytica
 N:Alternate names: ORF1
 C:Species: Pasteurella haemolytica
 C>Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Aug-1999
 C:Accession: JN0751
 R:Murphy, G.L.; Whitworth, L.C.
 Gene 129, 107-111, 1993
 A>Title: Analysis of tandem, multiple genes encoding 30-kDa membrane proteins in Pasteur
 A:Reference number: JN0751; MUID:93328110; PMID:8335249
 A:Accession: JN0751
 A:Molecule type: DNA
 A:Residues: 1-277 <NMR>
 A:Cross-references: GB:L11037; NID:g349529; PIDN:AAA25538.1; PID:g349530
 A:Experimental source: serotype A1
 A>Note: This protein displays a high degree of identity with an Escherichia coli inner m
 C:Comment: This protein is important in eliciting immunity to pneumonic pasteurellosis.
 C:Superfamily: lipoprotein-28
 C:Keywords: membrane protein

Query Match 12.3%; Score 34; DB 2; Length 277;
 Best Local Similarity 100.0%; Pred. No. 8e-26;
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 IAVNDPSNLAALILLEKGLIKDKNTNLFST 173
 DB 141 IAVNDPSNLAALILLEKGLIKDKNTNLFST 174

RESULT 2

Outer membrane protein, 28K - Haemophilus influenzae
 C:Species: Haemophilus influenzae
 C>Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 20-Aug-1999
 C:Accession: B64082; A43581
 R:Flaetschmann, R.D.; Adams, M.D.; White, O.; Cleyton, R.A.; Kirtness, E.F.; Kerlavage, A
 ; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, U
 ; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhmann, J.L.; Geoghagen, N.S.W.
 Science 269, 496-512, 1995
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
 A>Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
 A:Reference number: A64000; MUID:95350630; PMID:7542800
 A:Accession: B64082
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA

A:Residues: 1-273 <TIGR>
 A:Cross-references: GB:U32744; GB:L442023; NID:g1573608; PIDN:AAC22279.1; PID:g1573614; T
 A:Experimental source: Strain Rd KM20
 R:Chanyangam, M.; Smith, A.L.; Moseley, S.L.; Kuehn, M.; Jenny, P.
 Infect. Immun. 59, 600-608, 1991
 A>Title: Contribution of a 28-kilodalton membrane protein to the virulence of Haemophilus
 A:Reference number: A43581; MUID:91100034; PMID:1987077
 A:Accession: A43581
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 21-248; 'V', 250-273 <CHN>
 A:Cross-references: GB:M59804
 C:Superfamily: lipoprotein-28
 C:Keywords: membrane protein

Query Match 6.9%; Score 19; DB 2; Length 273;
 Best Local Similarity 100.0%; Pred. No. 7.6e-11;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 LNNLIVGNTFVYPLAGYS 125
 DB 104 LNNLIVGNTFVYPLAGYS 122

RESULT 3

Outer membrane 29.1K protein - Pasteurella haemolytica
 N:Alternate names: ORF3
 C:Species: Pasteurella haemolytica
 C>Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Aug-1999
 C:Accession: JN0753
 R:Murphy, G.L.; Whitworth, L.C.
 Gene 129, 107-111, 1993
 A>Title: Analysis of tandem, multiple genes encoding 30-kDa membrane proteins in Pasteur
 A:Reference number: JN0751; MUID:93328110; PMID:8335249
 A:Accession: JN0753
 A:Molecule type: DNA
 A:Residues: 1-263 <NMR>
 A:Cross-references: GB:L11037; NID:g349529; PIDN:AAA25540.1; PID:g349532
 A:Experimental source: serotype A1
 A>Note: This protein displays a high degree of identity with an Escherichia coli inner m
 C:Comment: This protein is important in eliciting immunity to pneumonic pasteurellosis.
 C:Superfamily: lipoprotein-28
 C:Keywords: membrane protein

Query Match 4.3%; Score 12; DB 2; Length 263;
 Best Local Similarity 100.0%; Pred. No. 0.00073;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 148 NLRALILLEKQ 159
 DB 135 NLRALILLEKQ 146

RESULT 4

Probable lipoprotein (imported) - Escherichia coli (strain O157:H7, substrain RIMD 050995)
 C:Species: Escherichia coli
 C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
 C:Accession: G90653
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 ; Gaeawa, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 Gena Res. 8, 11-22, 2001
 A>Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom
 A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: G90653
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-271 <HAV>
 A:Cross-references: GB:BA000007; PIDN:BAR33622.1; PID:g13359655; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain RIMD 0509952
 C:Genetics:
 A:Gene: Ecs0199

C:Superfamily: 1ipoprotein-28

Query Match 4.0%; Score 11; DB 2; Length 271;
Best Local Similarity 100.0%; Pred. No. 0.0074;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 224 FVEDKDSPPYV 234
Db 219 FVEDKDSPPYV 229

RESULT 5

G85504
C:Species: Escherichia coli (strain O157:H7, substrain EDL93)
C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C/Accession: G85504
R:Perla, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
11ter, L.; Grobbeck, E.U.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A/Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A/Reference number: A85480; MUID:21074935; PMID:11206551
A/Accession: G85504
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-271 <STO>
A/Cross-references: GB:AE005174; NID:912512928; PIDN:AA654499.1; GSPDB:GN00145; UWGP:Z02
A/Experimental source: strain O157:H7, substrain EDL933
C/Genetics:
A/Gene: YaeC
A:Superfamily: 1ipoprotein-28

Query Match 4.0%; Score 11; DB 2; Length 271;
Best Local Similarity 100.0%; Pred. No. 0.0074;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 224 FVEDKDSPPYV 234
Db 219 FVEDKDSPPYV 229

RESULT 6

A10532
C:Species: 1ipoprotein precursor STY0272 [imported] - Salmonella enterica subsp. enterica
C/Note: this species has also been called Salmonella typhi
A/Note: this species has also been called Salmonella typhi
C/Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C/Accession: A10532
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Crohlin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Garra, P.
Nature 413, 848-852, 2001
A/Authors: Park, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A/Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A/Reference number: AB0502; MUID:21594947; PMID:11677608
A/Accession: A10532
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-271 <PAR>
A/Cross-references: GB:AL513982; PIDN:CAD08705.1; PID:916501528; GSPDB:GN00176
C/Genetics:
A/Gene: STY0272
C:Superfamily: 1ipoprotein-28

Query Match 4.0%; Score 11; DB 2; Length 271;
Best Local Similarity 100.0%; Pred. No. 0.0074;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 224 FVEDKDSPPYV 234
Db 219 FVEDKDSPPYV 229

RESULT 7

A48107
A/Title: Gene RNA4 in Saccharomyces cerevisiae encodes the A12.2 subunit of RNA polymerase
A/Reference number: A48107; MUID:93109294; PMID:8417319
A/Accession: A48107
A/Molecule type: DNA
A/Residues: 1-125 <NOG>
A/Cross-references: EMBL:U00708; NID:9172461; PIDN:AAA34992.1; PID:9172462
A/Note: sequence extracted from NCBI backbone (NCBIN:121124, NCBI:121126)
A/Accession: B48107
A/Molecule type: DNA
A/Residues: 40-46/48-61/104-113 <NO2>
R:Baker, R.T.; Varshavsky, A.
submitted to the EMBL Data Library, September 1994
A/Description: N-terminal amidease: a new enzyme and component of a targeting complex in t

A/Reference number: S47937
A/Accession: S47937
A/Molecule type: DNA
A/Residues: 1-125 <BAK>
A/Cross-references: EMBL:U35664; NID:9531230; PIDN:AA659319.1; PID:9531231
R:Huang, M.E.; Chuat, J.C.; Galibert, F.
submitted to the Protein Sequence Database, September 1995
A/Reference number: S57052
A/Accession: S57052
A/Molecule type: DNA
A/Residues: 1-125 <MAN>
A/Cross-references: EMBL:U49563; NID:91015736; PIDN:CAA89591.1; PID:91015737; MIPS:YXR06;
R:Huang, M.E.; Manns, V.; Chuat, J.C.; Galibert, F.
Yeast 12, 869-875, 1996
A/Title: Analysis of a 62 kb DNA sequence of chromosome X reveals 36 open reading frames
A/Reference number: S71676; MUID:96437976; PMID:8840504
A/Accession: S71685
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA

A/Residues: 1-125 <HUA>
A/Cross-references: EMBL:U47993; NID:91019675; PIDN:AA639289.1; PID:91019665
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995
C/Genetics:
A/Gene: SGD:RPA12; RNA4
A/Cross-references: SGD:S0003824; MIPS:YXR063w
A/Map position: 10R
C:Superfamily: DNA-directed RNA polymerase chain A12.2/B12.6/ C11/M/14.5K
C/Keywords: nucleotidyltransferase; nucleus; transcription

Query Match 3.3%; Score 9; DB 2; Length 125;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 132 NELMDGATI 140
Db 74 NELMDGATI 82

RESULT 8

A1110
C:Species: 1ipoprotein lmo0285 [imported] - Listeria monocytogenes (strain EGD-e)
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C/Accession: A1110
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amand, A.; Baguer, F.; Berche, P.; Blocker,
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl, H.;
Science 294, 849-852, 2001

A:Reference number: AB1077; MWID:21537279; PMID:11679669
 A:Accession: AF110
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-273 <GUA>
 A:Cross-references: GB:NC_00310; PIDN:CAD00812.1; PID:G16409650; GSPDB:GN00177
 A:Experimental source: strain EGD-e
 C:Genetics:
 A:Gene: lmo0285
 C:Superfamily: lipoprotein-28

Query Match 3.3%; Score 9; DB 2; Length 273;
 Best Local Similarity 100.0%; Pred. No. 0.75;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 IAVPNDPSN 148
 |||||
 DB 136 IAVPNDPSN 144

RESULT 9
 A82266
 lipoprotein YaeC VC0905 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
 C:Species: Vibrio cholerae
 C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
 C:Accession: A82266
 R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
 Chaudson, D.; Ermolenko, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F.
 1, R.R.; Mekalanos, J.D.; Venter, U.C.; Fraser, C.M.
 Nature 406: 477-483, 2000
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A:Reference number: A82035; MWID:20406833; PMID:10952301
 A:Accession: A82266
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-275 <HEI>
 A:Cross-references: GB:AE004174; GB:AE003852; NID:G9655355; PIDN:AAF94067.1; GSPDB:GN001
 C:Experimental source: serogroup O1, strain N16961; biotype El Tor
 C:Genetics:
 A:Gene: VC0905
 A:Map position: 1
 C:Superfamily: lipoprotein-28

Query Match 3.3%; Score 9; DB 2; Length 275;
 Best Local Similarity 100.0%; Pred. No. 0.75;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 112 IVGNFVYP 120
 |||||
 DB 111 IVGNFVYP 119

RESULT 10
 A13551
 ABC transporter substrate binding protein BME10338 [imported] - Brucella melitensis (str
 C:Species: Brucella melitensis
 C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002
 C:Accession: A13551
 R:DelVecchio, V.G.; Kaparat, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
 .; Mazur, M.; Gelman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
 Proc. Natl. Acad. Sci. U.S.A. 99: 443-448, 2002
 A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
 A:Reference number: AD3252; PMID:11756688
 A:Accession: A13551
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-278 <KUP>
 A:Cross-references: GB:AE00918; PIDN:AAL53580.1; PID:G17984491; GSPDB:GN00191
 A:Experimental source: strain 16M
 C:Genetics:
 A:Gene: BME10338
 A:Map position: 11
 C:Superfamily: lipoprotein-28

Query Match 3.3%; Score 9; DB 2; Length 278;
 Best Local Similarity 100.0%; Pred. No. 0.76;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 DIVENPKXL 184
 |||||
 DB 177 DIVENPKXL 185

RESULT 11
 MH007
 Ig mu heavy chain disease protein (Bot) - human
 C:Species: Homo sapiens (man)
 C>Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 16-Jul-1999
 C:Accession: A02163
 R:Barikol-Matanabe, S.; Mihascu, E.; Mihascu, C.; Barikol, H.U.; Hilschmann, N.
 Hoppe-Seyler's Z. Physiol. Chem. 365: 105-118, 1984
 A:Title: The primary structure of mu-chain-disease protein BOT. Peculiar amino-acid sequ
 A:Reference number: A02163; MWID:84184186; PMID:6425189
 A:Accession: A02163
 A:Molecule type: Protein
 A:Residues: 1-391 <BAR>
 C:Comment: This protein has no V region homology or CH1 region.
 C:Genetics:
 A:Gene: GDB:IGHM
 A:Cross-references: GDB:120086; OMIM:147020
 A:Map position: 1432.33-1432.33
 C:Superfamily: Immunoglobulin C region, immunoglobulin homology
 C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin; transmembrane prot
 F:1-41/Domain: pre-C <VAR>
 F:43-391/Domain: Ig mu chain C region, secreted form <IGH>
 F:65-137/Domain: immunoglobulin homology <IGH1>
 F:175-243/Domain: immunoglobulin homology <IGH2>
 F:282-353/Domain: immunoglobulin homology <IGH3>
 F:447,210,217,378/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 3.3%; Score 9; DB 1; Length 391;
 Best Local Similarity 100.0%; Pred. No. 1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 TDEVEAEAK 265
 |||||
 DB 99 TDEVEAEAK 107

RESULT 12
 E81284
 hypothetical protein Cj1397 [imported] - Campylobacter jejuni (strain NCTC 11169)
 C:Species: Campylobacter jejuni
 C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
 C:Accession: E81284
 R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kelsey, G.M.; Churcher, C.; Basham, D.; Chilling
 C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell
 Nature 403: 665-668, 2000
 A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
 A:Reference number: A81250; MWID:20150912; PMID:10686204
 A:Accession: E81284
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-74 <PAR>
 A:Cross-references: GB:AL139078; GB:AL111168; NID:G9686723; PIDN:CAB73821.1; PID:G9686826
 A:Experimental source: serotype O2, strain NCTC 11168
 C:Genetics:
 A:Gene: Cj1395; Cj1397

Query Match 2.9%; Score 8; DB 2; Length 74;
 Best Local Similarity 100.0%; Pred. No. 2.4;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 TUNEKDG 137
 |||||
 DB 2 TUNEKDG 9

RESULT 13

molybdopterin converting factor, chain 1 (moad) PAB3357 - *Pyrococcus abyssi* (strain Orsa H75053)
C/Species: *Pyrococcus abyssi*
C/Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 17-Feb-2003
C/Accession: H75053
R:Anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A/Description: *Pyrococcus abyssi* genome sequence: insights into archaeal chromosome stru
A/Reference number: A75001
A/Accession: H75053
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-89 <KAM>
A/Cross-references: GB:AJ248287; GB:AL096836; NID:g5458657; PIDN:CAB50325.1; PID:e151622
A/Experimental source: strain Orsay
C/Genetics:
A/Gene: PAB3357
C/Superfamily: molybdopterin biosynthesis sulfur carrier protein
F/89/Modified site: 1-thioglycine (Gly) #status predicted

Query Match

Best Local Similarity 2.9%; Score 8; DB 2; Length 89;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 ELKDGATI 140

DB 23 ELKDGATI 30

RESULT 14

outer membrane lipoprotein Atu4489 [imported] - *Agrobacterium tumefaciens* (strain C58, D AE3108)
C/Species: *Agrobacterium tumefaciens*
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C/Accession: AE3108
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gallet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McChell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A/Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
A/Reference number: AB2577; MUID:21608550; PMID:11743193
A/Accession: AE3108
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-259 <KUP>
A/Cross-references: GB:AE008689; PIDN:AA45283.1; PID:g17742970; GSPDB:GN00187
A/Experimental source: strain C58 (Dupont)
C/Genetics:
A/Gene: Atu4489
A/Map position: linear chromosome
C/Superfamily: lipoprotein-28

Query Match

Best Local Similarity 2.9%; Score 8; DB 2; Length 259;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 ELKDGATI 140

DB 116 ELKDGATI 123

RESULT 15

probable periplasmic protein CJO772c [imported] - *Agrobacterium tumefaciens* (strain C58, G98178)
C/Species: *Agrobacterium tumefaciens*
C/Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 18-Nov-2002
C/Accession: G98178
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, W.; Qurollo, B.; Goldman,

A.; Liu, F.; Mollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tume*
A/Reference number: A97359; MUID:21608551; PMID:11743194

A/Accession: G98178
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-259 <KUP>

A/Cross-references: GB:AE007870; PIDN:AAK88953.1; PID:g15158732; GSPDB:GN00170
C/Genetics:
A/Gene: AGR_L_761
A/Map position: linear chromosome
C/Superfamily: lipoprotein-28

Query Match

Best Local Similarity 2.9%; Score 8; DB 2; Length 259;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 ELKDGATI 140

DB 116 ELKDGATI 123

RESULT 16

probable lipoprotein YaeC - *Escherichia coli* (strain K-12) E64744
C/Species: *Escherichia coli*
C/Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C/Accession: E64744; A47040
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A/Title: The complete genome sequence of *Escherichia coli* K-12.
A/Reference number: A64720; MUID:97426617; PMID:9278503
A/Accession: E64744
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-271 <BLAT>
A/Cross-references: GB:AE000129; GB:U00096; NID:g1786395; PIDN:AACT3308.1; PID:g1786396;
A/Experimental source: strain K-12, substrain MG1655
J.Gervais, F.G.; Drapeau, G.R.
J. Bacteriol. 174, 8016-8022, 1992
A/Title: Identification, cloning, and characterization of rcsF, a new regulator gene for
A/Reference number: A47040; MUID:93094132; PMID:1459591
A/Contents: K-12
A/Accession: A47040
A/Status: preliminary
A/Molecule type: nucleic acid
A/Residues: 127-199, 'L', 201-216, 'VSUKT', 223, 'SPRT' <GER>
A/Cross-references: GB:U04474; NID:g147530; PIDN:AA42507.1; PID:g147531
A/Note: sequence extracted from NCBI backbone (NCBIN:119949, NCBI:P.119950)
C/Genetics:
A/Gene: YaeC
C/Superfamily: lipoprotein-28

Query Match

Best Local Similarity 2.9%; Score 8; DB 2; Length 271;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 VGNTFYVP 120

DB 106 VGNTFYVP 113

RESULT 17

probable ATP-dependent RNA helicase PA3466 [imported] - *Pseudomonas aeruginosa* (strain P A83213)
C/Species: *Pseudomonas aeruginosa*
C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C/Accession: A83213
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bri
adman, S.; Yun, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lam,
.Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho
A:Reference number: A82950; PMID:20437337; PMID:10984043
A:Accession: A83213
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-446 <STO>
A:Cross-references: GB:AE004767; GB:AE004091; NID:g9949604; PIDN:AA06854.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA3466

Query Match 2.5%; Score 8; DB 2; Length 446;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 149 LAAALILL 156
|||||
Db 76 LAAALILL 83

RESULT 18
C81400
hypothetical protein Cj0539 [imported] - *Campylobacter jejuni* (strain NCTC 11168)
C:Species: *Campylobacter jejuni*
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
C:Accession: C81400
R:Hayashi, T.; Wren, B.W.; Mungall, K.; Kelsey, J.M.; Churcher, C.; Basham, D.; Chilling
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kelsey, J.M.; Churcher, C.; Basham, D.; Chilling
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; Vanylic, A.; Whitehead, S.; Barrer
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hyp
A:Reference number: A81250; PMID:20150912; PMID:10688204
A:Accession: C81400
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-87 <PAR>
A:Cross-references: GB:AL139075; GB:AL111168; NID:g6967817; PIDN:CA875175.1; PID:g696800
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: Cj0539

Query Match 2.5%; Score 7; DB 2; Length 87;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 155 LLEKQGL 161
|||||
Db 45 LLEKQGL 51

RESULT 19
B88029
protein F46F5.13 [imported] - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: B88029
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biolog
A:Reference number: A75000; PMID:99059613; PMID:9851916
A>Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C_ele
A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: B88029
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-104 <STO>
A:Cross-references: GB:chr_II; PIDN:AC78191.1; PID:g3886040; GSPDB:GN00020; CESP:F46F5.1
C:Genetics:
A:Gene: F46F5.13
A:Map position: 2

Query Match 2.5%; Score 7; DB 2; Length 104;
Best Local Similarity 100.0%; Pred. No. 32;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 108 NNVIIVG 114
|||||
Db 17 NNVIIVG 23

RESULT 20
C90820
hypothetical protein Ecs1531 [imported] - *Escherichia coli* (strain O157:H7, substrain R1
C:Species: *Escherichia coli*
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: C90820
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genom
A:Reference number: A99629; PMID:21156231; PMID:11258796
A:Accession: C90820
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-114 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA034954.1; PID:g13360995; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain R1MD 0509952
C:Genetics:
A:Gene: Ecs1531

Query Match 2.5%; Score 7; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 261 EBAKKQ 267
|||||
Db 54 EBAKKQ 60

RESULT 21
G30851
hypothetical protein Ecs1783 [imported] - *Escherichia coli* (strain O157:H7, substrain R1
C:Species: *Escherichia coli*
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: G30851
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genom
A:Reference number: A99629; PMID:21156231; PMID:11258796
A:Accession: G30851
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-114 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA035206.1; PID:g13361248; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain R1MD 0509952
C:Genetics:
A:Gene: Ecs1783

Query Match 2.5%; Score 7; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 261 EBAKKQ 267
|||||
Db 54 EBAKKQ 60

RESULT 22
C90874
hypothetical protein Ecs1963 [imported] - *Escherichia coli* (strain O157:H7, substrain R1
C:Species: *Escherichia coli*
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: C90874
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genc
A:Reference number: A99629; PMID:21156231; PMID:11258796
A:Accession: C90874
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-114 <HAY>
A:Cross-references: GB:BA000007; PIDN:BAE35386.1; PID:g13361428; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain R1MD 0509952
C:Genetics:
A:Gene: ECs1963

Query Match 2.5%; Score 7; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 261 EAAKKQ 267
DB 54 EAAKKQ 60

RESULT 23

hypothetical protein ECs2187 [imported] - *Escherichia coli* (strain O157:H7, substrain R1
C:Species: *Escherichia coli*
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: C90902
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genc
A:Reference number: A99629; PMID:21156231; PMID:11258796
A:Accession: C90902
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-114 <HAY>
A:Cross-references: GB:BA000007; PIDN:BAE35610.1; PID:g13361653; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain R1MD 0509952
C:Genetics:
A:Gene: ECs2187

Query Match 2.5%; Score 7; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 261 EAAKKQ 267
DB 54 EAAKKQ 60

RESULT 24

hypothetical protein ECs2260 [imported] - *Escherichia coli* (strain O157:H7, substrain R1
C:Species: *Escherichia coli*
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: D90911
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genc
A:Reference number: A99629; PMID:21156231; PMID:11258796
A:Accession: D90911
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-114 <HAY>
A:Cross-references: GB:BA000007; PIDN:BAE35683.1; PID:g13361726; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain R1MD 0509952
C:Genetics:
A:Gene: ECs2260

Query Match 2.5%; Score 7; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 261 EAAKKQ 267
DB 54 EAAKKQ 60

RESULT 25

hypothetical protein ECs2742 [imported] - *Escherichia coli* (strain O157:H7, substrain R1
C:Species: *Escherichia coli*
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: F90971
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genc
A:Reference number: A99629; PMID:21156231; PMID:11258796
A:Accession: F90971
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-114 <HAY>
A:Cross-references: GB:BA000007; PIDN:BAE36165.1; PID:g13362210; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain R1MD 0509952
C:Genetics:
A:Gene: ECs2742

Query Match 2.5%; Score 7; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 261 EAAKKQ 267
DB 54 EAAKKQ 60

RESULT 26

hypothetical protein ECs3496 [imported] - *Escherichia coli* (strain O157:H7, substrain R1
C:Species: *Escherichia coli*
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: H91065
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genc
A:Reference number: A99629; PMID:21156231; PMID:11258796
A:Accession: H91065
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-114 <HAY>
A:Cross-references: GB:BA000007; PIDN:BAE36919.1; PID:g13362967; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain R1MD 0509952
C:Genetics:
A:Gene: ECs3496

Query Match 2.5%; Score 7; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 261 EAAKKQ 267
DB 54 EAAKKQ 60

RESULT 27

unknown protein encoded by prophage CP-933N [imported] - *Escherichia coli* (strain O157:H7
C:Species: *Escherichia coli*
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: D85679
R:Perna, N.T.; Plunkett, III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,
Miller, L.; Grothbeck, C.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001

A>Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: D85679
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-114 <STO>
A:Cross-references: GB:AE005174; NID:G12514706; PIDN:AA655896.1; GSPDB:GN00145; UWGP:Z17
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: Z1795

Query Match 2.5%; Score 7; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 261 EBAKKQ 267
54 EBAKKQ 60

RESULT 28

unknown protein encoded within prophage CP-933U [imported] - *Escherichia coli* (strain O157:H7)
C:Species: *Escherichia coli*
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001

A:Accession: F85819
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-115 <STO>
A:Cross-references: GB:AE005174; NID:G12516122; PIDN:AA657018.1; GSPDB:GN00145; UWGP:Z31
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: Z3105

Query Match 2.5%; Score 7; DB 2; Length 115;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 261 EBAKKQ 267
55 EBAKKQ 61

RESULT 29

hypothetical protein Z1351 [imported] - *Escherichia coli* (strain O157:H7, substrain EDL933)
C:Species: *Escherichia coli*
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001

A:Accession: H85627; H85709
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-115 <STO>
A:Cross-references: GB:AE005174; NID:G12514192; PIDN:AA655464.1; GSPDB:GN00145; UWGP:Z13
A:Experimental source: strain O157:H7, substrain EDL933

A:Accession: H85709
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-115 <STO>
A:Cross-references: GB:AE005174; NID:G12515016; PIDN:AA656140.1; GSPDB:GN00145; UWGP:Z20
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:

Query Match 2.5%; Score 7; DB 2; Length 124;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 71 EBNDYAM 77
8 EBNDYAM 14

A:Gene: Z1351; Z2070
Query Match 2.5%; Score 7; DB 2; Length 115;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 261 EBAKKQ 267
55 EBAKKQ 61

RESULT 30
F85744
unknown protein encoded within prophage CP-933R [imported] - *Escherichia coli* (strain O157:H7)
C:Species: *Escherichia coli*
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001

A:Accession: F85744
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-123 <STO>
A:Cross-references: GB:AE005174; NID:G1251367; PIDN:AA656418.1; GSPDB:GN00145; UWGP:Z23
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: Z2372

Query Match 2.5%; Score 7; DB 2; Length 123;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 261 EBAKKQ 267
54 EBAKKQ 60

RESULT 31

hypothetical protein HP0712 - *Helicobacter pylori* (strain 26695)
C:Species: *Helicobacter pylori*
C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999

A:Accession: H64608
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-124 <TOM>
A:Cross-references: GB:AE00584; GB:AE00511; NID:G2313834; PIDN:AA07770.1; PID:G231384

Query Match 2.5%; Score 7; DB 2; Length 124;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 71 EBNDYAM 77
8 EBNDYAM 14

RESULT 32

A82559
integration host factor, beta subunit XF2437 [imported] - *Xylella fastidiosa* (strain 9a5)
C:Species: *Xylella fastidiosa*
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 17-Nov-2000

Query Match 2.5%; Score 7; DB 2; Length 124;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 71 EBNDYAM 77
8 EBNDYAM 14

C/Accession: A82559
R/Anonymous: The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A/Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A/Reference number: A82515; MUID:12035717; PMID:10910347
A/Note: for a complete list of authors see reference number A59328 below
A/Accession: A82559
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-126 <SIM>
A/Cross-references: GB:AE004052; GB:AE003849; NID:g105810617; PIND:AAFB5236.1; GSPDB:GN001
A/Experimental source: strain 95ac
R/Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briano, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carrer, D.M.; Carrex, H
as-Neco, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to Genbank, June 2000
A/Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigt
chado, W.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, E
A/Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmeri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A/Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsuchiko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A/Reference number: A59328
A/Content: annotation
C/Genetics:
A/Genes: XF2437
C/Superfamily: bacterial DNA-binding protein
Query Match 2.5%; Score 7; DB 2; Length 126;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 199 DDVDLAV 205
DB 42 DDVDLAV 48
RESULT 33
A84312
hypothetical protein Vng1589c [imported] - Halobacterium sp. NRC-1
C/Species: Halobacterium sp. NRC-1
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C/Accession: A84312
R/Ing, M.V.; Kennedy, S.P.; Mahataas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jbld
Jung, K.H.; Alam, M.; Freilich, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A/Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li
A/Title: Genome sequence of Halobacterium species NRC-1.
A/Reference number: A84160; MUID:20504483; PMID:11016950
A/Accession: A84312
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-137 <STO>
A/Cross-references: GB:AE004437; NID:g10581069; PIND:AAJ19861.1; GSPDB:GN00138
C/Genetics:
A/Genes: VNG1589c
Query Match 2.5%; Score 7; DB 2; Length 137;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 16 IALAGCS 22
DB 16 IALAGCS 22
RESULT 34
JE0359
plasmidocyte spreading peptide precursor protein - Pseudoplusia includens
C/Species: Pseudoplusia includens

C/Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 07-May-1999
C/Accession: JE0359
R/Clark, K.D.; Witherell, A.; Strand, M.R.
Biochem. Biophys. Res. Commun. 250, 479-485, 1998
A/Title: Plasmidocyte spreading peptide is encoded by an mRNA differentially expressed in
A/Reference number: JE0359; MUID:98440825; PMID:9753657
A/Accession: JE0359
A/Molecule type: mRNA
A/Residues: 1-141 <CLA>
C/Genetics:
A/Genes: p15
C/Superfamily: paralytic peptide I
Query Match 2.5%; Score 7; DB 2; Length 141;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 223 VFVEDKD 229
DB 65 VFVEDKD 71
RESULT 35
S74334
biotin carboxyl carrier protein - Synechocystis sp. (strain PCC 6803)
N/Alternate names: hypothetical protein slr0435
C/Species: Synechocystis sp.
A/Variety: PCC 6803
C/Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C/Accession: S74334
R/Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
o, K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda,
DNA Res. 3, 109-116, 1996
A/Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
S.
A/Reference number: S74322; MUID:97061201; PMID:8995231
A/Accession: S74334
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-154 <KAN>
A/Cross-references: EMBL:D64001; GB:AB001339; NID:g1001102; PIND:BA10252.1; PID:g100111;
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C/Genetics:
A/Genes: accB
A/Start codon: GTG
C/Superfamily: biotin carboxyl carrier protein, lipoyl/biotin-binding homology
F/73-153/Domain: lipoyl/biotin-binding homology <LPB>
Query Match 2.5%; Score 7; DB 2; Length 154;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 52 AEVAGQV 58
DB 126 AEVAGQV 132
RESULT 36
I67751
dopamine receptor D5 - green monkey (fragment)
C/Species: Cercopithecus aethiops (green monkey, grivet)
C/Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 13-Aug-1999
C/Accession: I67751
R/Nguyen, T.; Bard, J.A.; Jin, H.; Tarnesio, D.; Ward, D.C.; Kennedy, J.L.; Weinschenk, R.
Gene 109, 211-218, 1991
A/Title: Human dopamine D5 receptor pseudogenes.
A/Reference number: I53655; MUID:92112045; PMID:1765268
A/Accession: I67751
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-156 <R5>
A/Cross-references: GB:M77187; NID:g342376; PIND:AAA36923.1; PID:g342377
C/Genetics:

A:Gene: DRD5
C:Superfamily: vertebrate rhodopsin

Query Match 2.5%; Score 7; DB 2; Length 156;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 VAAVAG 56
|||
Db 25 VAAVAG 31

RESULT 37

C97438

hypothetical protein AGR_C.1160 [imported] - Agrobacterium tumefaciens (strain C58, Cere

C:Species: Agrobacterium tumefaciens

C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002

C:Accession: C97438

R:Goedert, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quirillo, B.; Goldman,

A.; Liu, F.; Woliam, C.; Allinger, W.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;

Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A:Reference number: A97359; PMID:21608551; PMID:11743194

A:Accession: C97438

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-188 <KUR>

A:Cross-references: GB:AE007869; PIDN:AAK86460.1; PID:G15155604; GSPDB:GN00169

C:Genetics:

A:Gene: AGR_C.1160

A:Map position: circular chromosome

Query Match 2.5%; Score 7; DB 2; Length 188;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 VAAARAD 199
|||
Db 71 VAAARAD 77

RESULT 38

AF2656

conserved hypothetical protein Atu0652 [imported] - Agrobacterium tumefaciens (strain C5

C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002

C:Accession: AF2656

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kuyavlin, T.; Levy, R.; Li, M.; McClell

; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, N.; Krespan, W.; Perry, M.; Gordon-Kamm,

ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; PMID:21608550; PMID:11743193

A:Accession: AF2656

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-188 <KUR>

A:Cross-references: GB:AE008688; PIDN:AA14166.1; PID:G17739011; GSPDB:GN00186

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: Atu0652

A:Map position: circular chromosome

Query Match 2.5%; Score 7; DB 2; Length 188;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 VAAARAD 199
|||
Db 71 VAAARAD 77

RESULT 39
JQ2185

coat protein - apple chlorotic leaf spot virus (isolate P-205)

N:Alternate names: ORF 3 protein

C:Species: apple chlorotic leaf spot virus

C:Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 20-Jun-2000

C:Accession: JQ2185

R:Sato, K.; Yoshikawa, N.; Takahashi, T.

J. Gen. Virol. 74, 1927-1931, 1993

A:Title: Complete nucleotide sequence of the genome of an apple isolate of apple chlorot

A:Reference number: JQ2183; PMID:93389448; PMID:8376968

A:Accession: JQ2185

A:Molecule type: mRNA

A:Residues: 1-193 <SKAT>

A:Cross-references: GB:D14966; NID:9434059; PIDN:BA03643.1; PID:9453242

C:Superfamily: apple chlorotic leaf spot virus coat protein

C:Keywords: coat protein

Query Match 2.5%; Score 7; DB 2; Length 193;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 TNLFSTT 174
|||
Db 126 TNLFSTT 132

RESULT 40

AC3629

3-oxoaprenyl-4-hydroxybenzoate carboxy-lyase (EC 4.1.1.-) [imported] - Brucella melitensis

C:Species: Brucella melitensis

C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002

C:Accession: AC3629

R:DeLVecchio, V.G.; Kapatal, V.; Redhar, R.J.; Patra, G.; Muier, C.; Los, T.; Ivanova, J

; Mazur, M.; Goldsman, E.; Selkov, E.; Blizer, P.H.; Hagluis, S.; O'Callaghan, D.; Letessc

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis

A:Reference number: AD3252; PMID:11756688

A:Accession: AC3629

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-196 <KUR>

A:Cross-references: GB:AE008918; PIDN:AA154198.1; PID:G17985166; GSPDB:GN00191

A:Experimental source: strain 16M

C:Genetics:

A:Gene: BMET10956

A:Map position: II

C:Superfamily: deaF protein

C:Keywords: carbon-carbon lyase; carboxy-lyase

Query Match 2.5%; Score 7; DB 2; Length 196;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 VAAARAD 199
|||
Db 167 VAAARAD 173

Search completed: June 16, 2004, 11:17:36
Job time : 32 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 16, 2004, 11:14:18 ; Search time 23 Seconds
(without alignments)
619.512 Million cell updates/sec

Title: US-10-018-672-2

Perfect score: 276

Sequence: 1 MNFGKINGICATASGIALG.....TDEVAEAKKQFDVIGKM 276

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 389414 seqs, 51625971 residues

Word size : 0

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

Issued Patents AA: *
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCUTS_COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57	20.7	118	US-09-540-236-2991	Sequence 2991, Ap
2	16	5.8	279	US-09-543-681A-6534	Sequence 6534, Ap
3	10	3.6	286	US-09-328-352-6702	Sequence 6702, Ap
4	9	3.3	285	US-09-489-039A-10641	Sequence 10641, A
5	9	3.3	291	US-09-107-532A-7234	Sequence 7234, Ap
6	9	3.3	296	US-09-328-352-6924	Sequence 6924, Ap
7	8	2.9	269	US-09-107-532A-6443	Sequence 6443, Ap
8	8	2.9	273	US-09-252-991A-19629	Sequence 19629, A
9	8	2.9	279	US-09-543-681A-4548	Sequence 4548, Ap
10	8	2.9	279	US-09-543-681A-6914	Sequence 6914, Ap
11	8	2.9	289	US-09-489-039A-12337	Sequence 12337, A
12	8	2.9	465	US-09-252-991A-17635	Sequence 17635, A
13	7	2.5	78	US-08-858-207A-422	Sequence 422, App
14	7	2.5	116	US-09-134-001C-3331	Sequence 3331, Ap
15	7	2.5	159	US-09-963-791-18	Sequence 18, Appl
16	7	2.5	161	US-08-572-447C-11	Sequence 11, Appl
17	7	2.5	161	US-09-267-747-11	Sequence 11, Appl
18	7	2.5	199	PCT-US94-02889-2	Sequence 2, Appl
19	7	2.5	226	US-08-572-447C-15	Sequence 15, Appl
20	7	2.5	226	US-09-267-747-15	Sequence 15, Appl
21	7	2.5	274	US-09-543-681A-4708	Sequence 4708, Ap
22	7	2.5	275	US-09-489-039A-7722	Sequence 7722, Ap
23	7	2.5	298	US-09-434-354-49	Sequence 49, Appl
24	7	2.5	298	US-09-434-354-48	Sequence 49, Appl
25	7	2.5	302	US-09-134-001C-3171	Sequence 3171, Ap
26	7	2.5	306	US-09-252-991A-18673	Sequence 18673, A
27	7	2.5	309	US-09-522-714-10	Sequence 10, Appl

28	7	2.5	310	US-09-963-791-8	Sequence 8, Appl
29	7	2.5	317	US-09-963-791-16	Sequence 16, Appl
30	7	2.5	326	US-09-489-039A-9512	Sequence 9512, Ap
31	7	2.5	342	US-09-540-236-2111	Sequence 2111, Ap
32	7	2.5	347	US-09-543-681A-5961	Sequence 5961, Ap
33	7	2.5	351	US-09-252-991A-30094	Sequence 30094, A
34	7	2.5	356	US-09-963-791-20	Sequence 20, Appl
35	7	2.5	361	US-09-543-681A-5390	Sequence 5390, Ap
36	7	2.5	419	US-08-305-172B-4	Sequence 4, Appl
37	7	2.5	438	US-09-963-791-22	Sequence 22, Appl
38	7	2.5	447	US-09-252-991A-20116	Sequence 20116, A
39	7	2.5	468	US-09-963-791-6	Sequence 6, Appl
40	7	2.5	475	US-07-686-591-4	Sequence 4, Appl
41	7	2.5	475	US-07-970-715-4	Sequence 4, Appl
42	7	2.5	477	US-07-791-936A-2	Sequence 2, Appl
43	7	2.5	477	US-08-383-781B-2	Sequence 2, Appl
44	7	2.5	477	US-07-969-267B-2	Sequence 2, Appl
45	7	2.5	477	US-07-969-267B-3	Sequence 3, Appl
46	7	2.5	477	US-09-168-510-2	Sequence 2, Appl
47	7	2.5	477	US-09-168-510-3	Sequence 3, Appl
48	7	2.5	506	US-09-134-001C-4383	Sequence 4383, Ap
49	7	2.5	507	US-09-963-791-10	Sequence 10, Appl
50	7	2.5	581	US-09-543-681A-4472	Sequence 4472, Ap
51	7	2.5	589	US-09-963-791-12	Sequence 12, Appl
52	7	2.5	652	US-08-987-119-2	Sequence 2, Appl
53	7	2.5	690	US-09-336-115C-22	Sequence 22, Appl
54	7	2.5	691	US-09-336-115C-12	Sequence 12, Appl
55	7	2.5	757	US-09-963-791-24	Sequence 24, Appl
56	7	2.5	859	US-09-369-364A-5	Sequence 5, Appl
57	7	2.5	886	US-09-252-991A-26999	Sequence 26999, A
58	7	2.5	908	US-09-963-791-2	Sequence 2, Appl
59	7	2.5	982	US-09-252-991A-30580	Sequence 30580, A
60	7	2.5	1482	US-09-410-551B-21	Sequence 21, Appl
61	7	2.5	1488	US-09-410-551B-17	Sequence 17, Appl
62	7	2.5	1509	US-09-410-551B-23	Sequence 23, Appl
63	7	2.5	1517	US-09-410-551B-19	Sequence 19, Appl
64	7	2.5	1969	US-09-418-710-72	Sequence 72, Appl
65	7	2.5	1972	US-09-418-710-21	Sequence 21, Appl
66	7	2.5	6396	US-09-410-551B-72	Sequence 72, Appl
67	6	2.2	14	US-08-165-038-5	Sequence 5, Appl
68	6	2.2	14	US-08-876-781-5	Sequence 5, Appl
69	6	2.2	14	US-09-567-003C-2	Sequence 2, Appl
70	6	2.2	17	US-08-290-448A-55	Sequence 55, Appl
71	6	2.2	17	US-08-290-448A-55	Sequence 55, Appl
72	6	2.2	17	US-08-175-069A-55	Sequence 55, Appl
73	6	2.2	17	US-08-461-939B-55	Sequence 55, Appl
74	6	2.2	17	US-08-464-000-55	Sequence 55, Appl
75	6	2.2	20	US-09-009-993-193	Sequence 193, App
76	6	2.2	25	US-09-515-356-3	Sequence 3, Appl
77	6	2.2	25	US-09-515-356-4	Sequence 4, Appl
78	6	2.2	27	US-09-053-197A-69	Sequence 69, Appl
79	6	2.2	27	US-09-085-761A-74	Sequence 74, Appl
80	6	2.2	27	US-09-017-689A-1	Sequence 1, Appl
81	6	2.2	30	US-08-290-448A-17	Sequence 17, Appl
82	6	2.2	30	US-08-290-448A-17	Sequence 17, Appl
83	6	2.2	30	US-08-175-069A-17	Sequence 17, Appl
84	6	2.2	30	US-08-461-939B-17	Sequence 17, Appl
85	6	2.2	30	US-08-464-000-17	Sequence 17, Appl
86	6	2.2	34	US-08-129-089-5	Sequence 5, Appl
87	6	2.2	34	US-08-129-089-8	Sequence 8, Appl
88	6	2.2	34	US-08-129-089-9	Sequence 9, Appl
89	6	2.2	34	US-08-129-089-10	Sequence 10, Appl
90	6	2.2	34	US-08-129-089-11	Sequence 11, Appl
91	6	2.2	34	US-08-129-089-12	Sequence 12, Appl
92	6	2.2	34	US-08-454-444-1	Sequence 1, Appl
93	6	2.2	34	US-08-454-444-4	Sequence 4, Appl
94	6	2.2	34	US-08-454-444-5	Sequence 5, Appl
95	6	2.2	34	US-08-454-444-6	Sequence 6, Appl
96	6	2.2	34	US-08-454-444-7	Sequence 7, Appl
97	6	2.2	34	US-08-454-444-8	Sequence 8, Appl
98	6	2.2	34	PCT-US93-01135-5	Sequence 5, Appl
99	6	2.2	34	PCT-US93-01135-8	Sequence 8, Appl
100	6	2.2	34	PCT-US93-01135-9	Sequence 9, Appl

101	6	2.2	34	5	PCT-US93-01135-10	Sequence 10, Appl	174	6	2.2	166	4	US-09-339-913B-85	Sequence 85, Appl
102	6	2.2	34	5	PCT-US93-01135-11	Sequence 11, Appl	175	6	2.2	166	4	US-09-339-904A-75	Sequence 75, Appl
103	6	2.2	34	5	PCT-US93-01135-12	Sequence 12, Appl	176	6	2.2	166	4	US-09-339-904A-76	Sequence 76, Appl
104	6	2.2	42	1	US-08-290-448A-13	Sequence 13, Appl	177	6	2.2	166	4	US-09-339-904A-77	Sequence 77, Appl
105	6	2.2	42	1	US-08-290-448A-13	Sequence 13, Appl	178	6	2.2	166	4	US-09-339-904A-81	Sequence 81, Appl
106	6	2.2	42	1	US-08-175-069A-13	Sequence 13, Appl	179	6	2.2	166	4	US-09-339-904A-82	Sequence 82, Appl
107	6	2.2	42	1	US-08-461-939B-13	Sequence 13, Appl	180	6	2.2	166	4	US-09-339-904A-84	Sequence 84, Appl
108	6	2.2	42	4	US-08-461-939B-13	Sequence 13, Appl	181	6	2.2	166	4	US-09-339-904A-85	Sequence 85, Appl
109	6	2.2	44	4	US-08-858-207A-280	Sequence 280, Appl	182	6	2.2	166	4	US-09-339-904A-85	Sequence 85, Appl
110	6	2.2	44	1	US-08-290-448A-12	Sequence 12, Appl	183	6	2.2	166	4	US-09-339-904A-86	Sequence 86, Appl
111	6	2.2	45	1	US-08-290-448A-12	Sequence 12, Appl	184	6	2.2	166	4	US-09-339-904A-86	Sequence 86, Appl
112	6	2.2	45	1	US-08-175-069A-12	Sequence 12, Appl	185	6	2.2	166	4	US-08-769-062B-75	Sequence 75, Appl
113	6	2.2	45	4	US-08-461-939B-12	Sequence 12, Appl	186	6	2.2	166	4	US-08-769-062B-76	Sequence 76, Appl
114	6	2.2	45	4	US-08-461-939B-12	Sequence 12, Appl	187	6	2.2	166	4	US-08-769-062B-77	Sequence 77, Appl
115	6	2.2	54	1	US-08-165-038-1	Sequence 1, Appl	188	6	2.2	166	4	US-08-769-062B-79	Sequence 79, Appl
116	6	2.2	54	1	US-08-621-081A-1	Sequence 1, Appl	189	6	2.2	166	4	US-08-769-062B-81	Sequence 81, Appl
117	6	2.2	54	2	US-08-876-781-1	Sequence 1, Appl	190	6	2.2	166	4	US-08-769-062B-82	Sequence 82, Appl
118	6	2.2	54	4	US-08-936-165A-320	Sequence 320, Appl	191	6	2.2	166	4	US-08-769-062B-84	Sequence 84, Appl
119	6	2.2	58	4	US-09-621-976-662	Sequence 662, Appl	192	6	2.2	166	4	US-08-769-062B-85	Sequence 85, Appl
120	6	2.2	61	4	US-08-858-207A-343	Sequence 343, Appl	193	6	2.2	166	4	US-08-769-062B-86	Sequence 86, Appl
121	6	2.2	63	4	US-09-252-991A-19534	Sequence 19534, A	194	6	2.2	166	4	US-09-344-002B-75	Sequence 75, Appl
122	6	2.2	66	4	US-09-227-357-654	Sequence 654, Appl	195	6	2.2	166	4	US-09-344-002B-77	Sequence 77, Appl
123	6	2.2	70	4	US-09-489-039A-12939	Sequence 12939, A	196	6	2.2	166	4	US-09-344-002B-79	Sequence 79, Appl
124	6	2.2	82	4	US-09-252-991A-18536	Sequence 18536, A	197	6	2.2	166	4	US-09-344-002B-81	Sequence 81, Appl
125	6	2.2	82	4	US-09-489-039A-14014	Sequence 14014, A	198	6	2.2	166	4	US-09-344-002B-82	Sequence 82, Appl
126	6	2.2	84	4	US-09-252-991A-18487	Sequence 18487, A	199	6	2.2	166	4	US-09-344-002B-84	Sequence 84, Appl
127	6	2.2	92	4	US-09-621-976-5000	Sequence 5000, Ap	200	6	2.2	166	4	US-09-344-002B-85	Sequence 85, Appl
128	6	2.2	99	4	US-09-540-236-2992	Sequence 2992, Ap	201	6	2.2	166	4	US-09-344-002B-86	Sequence 86, Appl
129	6	2.2	100	4	US-09-732-210-109	Sequence 709, Appl	202	6	2.2	166	4	US-09-205-461-109	Sequence 109, Appl
130	6	2.2	107	4	US-09-328-352-5562	Sequence 5562, Ap	203	6	2.2	166	4	US-09-559-565C-75	Sequence 75, Appl
131	6	2.2	111	1	US-08-466-886-38	Sequence 38, Appl	204	6	2.2	166	4	US-09-559-565C-76	Sequence 76, Appl
132	6	2.2	111	3	US-08-469-617-38	Sequence 38, Appl	205	6	2.2	166	4	US-09-559-565C-77	Sequence 77, Appl
133	6	2.2	114	4	US-09-489-039A-8548	Sequence 8548, Ap	206	6	2.2	166	4	US-09-559-565C-79	Sequence 79, Appl
134	6	2.2	116	4	US-09-732-210-1691	Sequence 1691, Ap	207	6	2.2	166	4	US-09-559-565C-81	Sequence 81, Appl
135	6	2.2	117	1	US-07-942-245-24	Sequence 24, Appl	208	6	2.2	166	4	US-09-559-565C-82	Sequence 82, Appl
136	6	2.2	117	4	US-09-732-210-732	Sequence 732, Appl	209	6	2.2	166	4	US-09-559-565C-84	Sequence 84, Appl
137	6	2.2	118	4	US-09-732-210-147	Sequence 147, Appl	210	6	2.2	166	4	US-09-559-565C-85	Sequence 85, Appl
138	6	2.2	119	4	US-09-252-991A-29926	Sequence 29926, A	211	6	2.2	166	4	US-09-559-565C-86	Sequence 86, Appl
139	6	2.2	122	4	US-09-713-974B-349	Sequence 685, Appl	212	6	2.2	166	4	US-09-559-565C-86	Sequence 86, Appl
140	6	2.2	124	4	US-08-532-914A-2188	Sequence 349, Ap	213	6	2.2	166	4	US-09-252-991A-32785	Sequence 32785, A
141	6	2.2	127	4	US-09-252-991A-2188	Sequence 2188, A	214	6	2.2	166	4	US-09-693-350-75	Sequence 75, Appl
142	6	2.2	128	4	US-09-621-976-5844	Sequence 5844, Ap	215	6	2.2	166	4	US-09-693-350-76	Sequence 76, Appl
143	6	2.2	131	4	US-09-247-155-175	Sequence 175, Appl	216	6	2.2	166	4	US-09-693-350-77	Sequence 77, Appl
144	6	2.2	140	4	US-09-134-000C-6522	Sequence 6522, Ap	217	6	2.2	166	4	US-09-693-350-79	Sequence 79, Appl
145	6	2.2	141	4	US-09-107-532A-4712	Sequence 4712, Ap	218	6	2.2	166	4	US-09-693-350-81	Sequence 81, Appl
146	6	2.2	146	4	US-09-543-681A-7317	Sequence 7317, Ap	219	6	2.2	166	4	US-09-693-350-82	Sequence 82, Appl
147	6	2.2	147	4	US-09-134-001C-3024	Sequence 3024, Ap	220	6	2.2	166	4	US-09-693-350-84	Sequence 84, Appl
148	6	2.2	151	4	US-09-732-210-1674	Sequence 1674, Ap	221	6	2.2	166	4	US-09-693-350-85	Sequence 85, Appl
149	6	2.2	151	4	US-09-489-039A-9773	Sequence 9773, Ap	222	6	2.2	166	4	US-09-693-350-86	Sequence 86, Appl
150	6	2.2	152	4	US-09-252-991A-26218	Sequence 26218, A	223	6	2.2	166	4	US-09-693-389-75	Sequence 75, Appl
151	6	2.2	153	4	US-09-732-210-1672	Sequence 1672, Ap	224	6	2.2	166	4	US-09-693-389-76	Sequence 76, Appl
152	6	2.2	155	4	US-09-370-838-114	Sequence 114, Appl	225	6	2.2	166	4	US-09-693-389-77	Sequence 77, Appl
153	6	2.2	156	1	PCT-US94-06447-5	Sequence 5, Appl	226	6	2.2	166	4	US-09-693-389-79	Sequence 79, Appl
154	6	2.2	156	5	PCT-US94-06447-5	Sequence 5, Appl	227	6	2.2	166	4	US-09-693-389-81	Sequence 81, Appl
155	6	2.2	160	4	US-09-252-991A-25771	Sequence 25771, A	228	6	2.2	166	4	US-09-693-389-82	Sequence 82, Appl
156	6	2.2	161	4	US-09-252-991A-23802	Sequence 23802, A	229	6	2.2	166	4	US-09-693-389-84	Sequence 84, Appl
157	6	2.2	166	1	US-08-288-671-1	Sequence 1, Appl	230	6	2.2	166	4	US-09-693-389-85	Sequence 85, Appl
158	6	2.2	166	2	US-08-489-066A-16	Sequence 16, Appl	231	6	2.2	166	4	US-06-256-204C-72	Sequence 72, Appl
159	6	2.2	166	2	US-08-997-080-109	Sequence 109, Appl	232	6	2.2	166	4	US-06-256-204C-73	Sequence 73, Appl
160	6	2.2	166	2	US-08-875-802-4	Sequence 4, Appl	233	6	2.2	166	4	US-06-256-204C-75	Sequence 75, Appl
161	6	2.2	166	2	US-08-997-362-109	Sequence 109, Appl	234	6	2.2	166	4	US-06-256-204C-77	Sequence 77, Appl
162	6	2.2	166	3	US-08-489-072A-16	Sequence 16, Appl	235	6	2.2	166	4	US-06-256-204C-78	Sequence 78, Appl
163	6	2.2	166	3	US-09-095-855-109	Sequence 109, Appl	236	6	2.2	166	4	US-06-256-204C-80	Sequence 80, Appl
164	6	2.2	166	3	US-09-305-723C-4	Sequence 4, Appl	237	6	2.2	166	4	US-06-256-204C-83	Sequence 83, Appl
165	6	2.2	166	4	US-08-489-071A-16	Sequence 16, Appl	238	6	2.2	166	4	US-06-256-204C-85	Sequence 85, Appl
166	6	2.2	166	4	US-09-339-913B-75	Sequence 75, Appl	239	6	2.2	166	4	US-09-559-671A-75	Sequence 75, Appl
167	6	2.2	166	4	US-09-339-913B-76	Sequence 76, Appl	240	6	2.2	166	4	US-09-559-671A-76	Sequence 76, Appl
168	6	2.2	166	4	US-09-339-913B-77	Sequence 77, Appl	241	6	2.2	166	4	US-09-559-671A-77	Sequence 77, Appl
169	6	2.2	166	4	US-09-339-913B-79	Sequence 79, Appl	242	6	2.2	166	4	US-09-559-671A-79	Sequence 79, Appl
170	6	2.2	166	4	US-09-339-913B-81	Sequence 81, Appl	243	6	2.2	166	4	US-09-559-671A-81	Sequence 81, Appl
171	6	2.2	166	4	US-09-339-913B-82	Sequence 82, Appl	244	6	2.2	166	4	US-09-559-671A-81	Sequence 81, Appl
172	6	2.2	166	4	US-09-339-913B-84	Sequence 84, Appl	245	6	2.2	166	4	US-09-559-671A-82	Sequence 82, Appl
173	6	2.2	166	4	US-09-339-913B-85	Sequence 85, Appl	246	6	2.2	166	4	US-09-559-671A-84	Sequence 84, Appl

247	6	2.2	166	4	US-09-559-671A-85	Sequence 85, Appl	320	6	2.2	189	3	US-09-206-935-13	Sequence 13, Appl
248	6	2.2	166	4	US-09-559-671A-86	Sequence 86, Appl	321	6	2.2	189	3	US-09-206-935-14	Sequence 14, Appl
249	6	2.2	166	4	US-09-339-926A-75	Sequence 75, Appl	322	6	2.2	189	3	US-09-206-935-15	Sequence 15, Appl
250	6	2.2	166	4	US-09-339-926A-76	Sequence 76, Appl	323	6	2.2	189	3	US-09-206-935-17	Sequence 17, Appl
251	6	2.2	166	4	US-09-339-926A-77	Sequence 77, Appl	324	6	2.2	189	3	US-09-206-935-18	Sequence 18, Appl
252	6	2.2	166	4	US-09-339-926A-79	Sequence 79, Appl	325	6	2.2	189	3	US-09-206-935-19	Sequence 19, Appl
253	6	2.2	166	4	US-09-339-926A-81	Sequence 81, Appl	326	6	2.2	189	4	US-08-489-072A-2	Sequence 2, Appl
254	6	2.2	166	4	US-09-339-926A-82	Sequence 82, Appl	327	6	2.2	189	4	US-08-489-071A-3	Sequence 3, Appl
255	6	2.2	166	4	US-09-339-926A-84	Sequence 84, Appl	328	6	2.2	189	4	US-09-206-935-10	Sequence 10, Appl
256	6	2.2	166	4	US-09-339-926A-85	Sequence 85, Appl	329	6	2.2	189	4	US-09-206-935-13	Sequence 13, Appl
257	6	2.2	166	4	US-09-339-926A-86	Sequence 86, Appl	330	6	2.2	189	4	US-09-206-935-14	Sequence 14, Appl
258	6	2.2	166	4	US-09-744-754C-9	Sequence 9, Appl	331	6	2.2	189	4	US-09-206-935-15	Sequence 15, Appl
259	6	2.2	166	4	US-09-744-754C-11	Sequence 11, Appl	332	6	2.2	189	4	US-09-206-935-17	Sequence 17, Appl
260	6	2.2	166	4	US-09-744-754C-30	Sequence 30, Appl	333	6	2.2	189	4	US-09-206-935-18	Sequence 18, Appl
261	6	2.2	166	4	US-09-744-754C-32	Sequence 32, Appl	334	6	2.2	189	4	US-09-206-935-19	Sequence 19, Appl
262	6	2.2	166	4	US-09-744-754C-34	Sequence 34, Appl	335	6	2.2	189	4	US-09-487-793-7	Sequence 7, Appl
263	6	2.2	166	4	US-09-744-754C-36	Sequence 36, Appl	336	6	2.2	189	4	US-09-487-793-8	Sequence 8, Appl
264	6	2.2	166	4	US-09-744-754C-38	Sequence 38, Appl	337	6	2.2	189	4	US-09-598-594-7	Sequence 7, Appl
265	6	2.2	166	4	US-09-744-754C-40	Sequence 40, Appl	338	6	2.2	189	4	US-09-598-594-8	Sequence 8, Appl
266	6	2.2	166	4	US-09-744-754C-42	Sequence 42, Appl	339	6	2.2	189	4	US-07-145-002B-4	Sequence 4, Appl
267	6	2.2	167	2	US-08-875-802-2	Sequence 2, Appl	340	6	2.2	189	4	US-07-145-002B-6	Sequence 6, Appl
268	6	2.2	167	2	US-08-875-802-3	Sequence 3, Appl	341	6	2.2	189	4	US-07-145-002B-18	Sequence 18, Appl
269	6	2.2	167	3	US-09-305-723C-2	Sequence 2, Appl	342	6	2.2	189	4	US-07-145-002B-19	Sequence 19, Appl
270	6	2.2	167	3	US-09-305-723C-3	Sequence 3, Appl	343	6	2.2	189	4	US-07-145-002B-18	Sequence 18, Appl
271	6	2.2	167	4	US-07-145-002B-46	Sequence 46, Appl	344	6	2.2	189	4	US-07-145-002B-22	Sequence 22, Appl
272	6	2.2	167	4	US-07-145-002B-47	Sequence 47, Appl	345	6	2.2	189	4	US-07-145-002B-30	Sequence 30, Appl
273	6	2.2	167	4	US-07-145-002B-49	Sequence 49, Appl	346	6	2.2	189	4	US-07-145-002B-32	Sequence 32, Appl
274	6	2.2	167	4	US-07-145-002B-51	Sequence 51, Appl	347	6	2.2	189	4	US-07-145-002B-37	Sequence 37, Appl
275	6	2.2	167	4	US-07-145-002B-52	Sequence 52, Appl	348	6	2.2	189	4	US-06-256-204C-4	Sequence 4, Appl
276	6	2.2	167	4	US-07-145-002B-54	Sequence 54, Appl	349	6	2.2	189	4	US-06-256-204C-6	Sequence 6, Appl
277	6	2.2	167	4	US-07-145-002B-57	Sequence 57, Appl	350	6	2.2	189	4	US-06-256-204C-18	Sequence 18, Appl
278	6	2.2	167	4	US-07-145-002B-59	Sequence 59, Appl	351	6	2.2	189	4	US-06-256-204C-19	Sequence 19, Appl
279	6	2.2	167	4	US-06-256-204C-46	Sequence 46, Appl	352	6	2.2	189	4	US-06-256-204C-22	Sequence 22, Appl
280	6	2.2	167	4	US-06-256-204C-47	Sequence 47, Appl	353	6	2.2	189	4	US-06-256-204C-30	Sequence 30, Appl
281	6	2.2	167	4	US-06-256-204C-49	Sequence 49, Appl	354	6	2.2	189	4	US-06-256-204C-32	Sequence 32, Appl
282	6	2.2	167	4	US-06-256-204C-51	Sequence 51, Appl	355	6	2.2	189	4	US-06-256-204C-37	Sequence 37, Appl
283	6	2.2	167	4	US-06-256-204C-52	Sequence 52, Appl	356	6	2.2	189	4	US-09-352-991A-26824	Sequence 26824, A
284	6	2.2	167	4	US-06-256-204C-54	Sequence 54, Appl	357	6	2.2	203	4	US-08-956-171E-5207	Sequence 5207, A
285	6	2.2	167	4	US-06-256-204C-57	Sequence 57, Appl	358	6	2.2	203	4	US-08-956-171E-5205	Sequence 5205, A
286	6	2.2	167	4	US-06-256-204C-59	Sequence 59, Appl	359	6	2.2	203	4	US-09-134-000C-2975	Sequence 2975, A
287	6	2.2	169	4	US-07-145-002B-34	Sequence 34, Appl	360	6	2.2	205	4	US-09-252-991A-24225	Sequence 24225, A
288	6	2.2	169	4	US-07-145-002B-38	Sequence 38, Appl	361	6	2.2	209	4	US-09-134-000C-4884	Sequence 4884, A
289	6	2.2	169	4	US-06-256-204C-34	Sequence 34, Appl	362	6	2.2	210	4	US-09-543-681A-815	Sequence 815, A
290	6	2.2	169	4	US-06-256-204C-38	Sequence 38, Appl	363	6	2.2	210	4	US-09-328-038-4	Sequence 4, Appl
291	6	2.2	169	4	US-06-256-204C-58	Sequence 58, Appl	364	6	2.2	212	4	US-08-165-058-4	Sequence 4, Appl
292	6	2.2	170	4	US-09-134-000C-3567	Sequence 3567, A	365	6	2.2	213	1	US-08-621-081A-4	Sequence 4, Appl
293	6	2.2	170	4	US-09-252-991A-27069	Sequence 27069, A	366	6	2.2	213	1	US-08-621-081A-26	Sequence 26, Appl
294	6	2.2	172	4	US-09-328-352-6306	Sequence 6306, A	367	6	2.2	213	2	US-08-621-081A-27	Sequence 27, Appl
295	6	2.2	175	4	US-09-651-941-13	Sequence 13, Appl	368	6	2.2	213	2	US-08-621-081A-28	Sequence 28, Appl
296	6	2.2	175	4	US-09-651-941-13	Sequence 13, Appl	369	6	2.2	213	4	US-09-252-991A-28039	Sequence 28039, A
297	6	2.2	179	4	US-09-489-039A-8913	Sequence 8913, A	370	6	2.2	214	4	US-09-352-991A-19712	Sequence 19712, A
298	6	2.2	182	4	US-09-107-532A-6641	Sequence 6641, A	371	6	2.2	215	4	US-08-914-3755-29	Sequence 29, Appl
299	6	2.2	182	4	US-09-107-532A-6641	Sequence 6641, A	372	6	2.2	215	4	US-08-914-3755-29	Sequence 29, Appl
300	6	2.2	185	4	US-09-050-739-12	Sequence 12, Appl	373	6	2.2	215	4	US-09-124-884-6	Sequence 6, Appl
301	6	2.2	186	4	US-09-198-452A-106	Sequence 106, Appl	374	6	2.2	215	4	US-09-124-884-6	Sequence 6, Appl
302	6	2.2	186	4	US-09-489-039A-12559	Sequence 12559, A	375	6	2.2	217	1	US-08-621-081A-13	Sequence 13, Appl
303	6	2.2	189	6	5510472-8	Patent No. 5510472	376	6	2.2	217	1	US-08-621-081A-21	Sequence 21, Appl
304	6	2.2	189	1	US-08-026-758-1	Sequence 1, Appl	377	6	2.2	217	1	US-08-621-081A-22	Sequence 22, Appl
305	6	2.2	189	1	US-08-026-758-9	Sequence 9, Appl	378	6	2.2	217	1	US-08-621-081A-23	Sequence 23, Appl
306	6	2.2	189	1	US-08-026-758-10	Sequence 10, Appl	379	6	2.2	217	1	US-08-621-081A-24	Sequence 24, Appl
307	6	2.2	189	1	US-08-026-758-11	Sequence 11, Appl	380	6	2.2	217	1	US-08-621-081A-25	Sequence 25, Appl
308	6	2.2	189	1	US-08-026-758-12	Sequence 12, Appl	381	6	2.2	217	1	US-08-621-081A-26	Sequence 26, Appl
309	6	2.2	189	1	US-08-026-758-13	Sequence 13, Appl	382	6	2.2	217	1	US-08-621-081A-27	Sequence 27, Appl
310	6	2.2	189	1	US-08-026-758-14	Sequence 14, Appl	383	6	2.2	217	1	US-08-621-081A-28	Sequence 28, Appl
311	6	2.2	189	1	US-08-026-758-15	Sequence 15, Appl	384	6	2.2	217	1	US-08-621-081A-29	Sequence 29, Appl
312	6	2.2	189	1	US-08-026-758-16	Sequence 16, Appl	385	6	2.2	217	4	US-09-252-991A-24723	Sequence 24723, A
313	6	2.2	189	1	US-08-026-758-17	Sequence 17, Appl	386	6	2.2	219	1	US-08-266-451B-24	Sequence 24, Appl
314	6	2.2	189	1	US-08-026-758-18	Sequence 18, Appl	387	6	2.2	219	1	US-08-621-081A-16	Sequence 16, Appl
315	6	2.2	189	2	US-08-489-065A-2	Sequence 2, Appl	388	6	2.2	221	1	US-08-748-725-24	Sequence 24, Appl
316	6	2.2	189	2	US-08-489-066A-3	Sequence 3, Appl	389	6	2.2	221	1	US-08-621-081A-19	Sequence 19, Appl
317	6	2.2	189	3	US-08-489-072A-2	Sequence 2, Appl	390	6	2.2	221	4	US-09-252-991A-22003	Sequence 22003, A
318	6	2.2	189	3	US-08-489-072A-3	Sequence 3, Appl	391	6	2.2	221	4	US-09-252-991A-32061	Sequence 32061, A
319	6	2.2	189	3	US-09-206-935-10	Sequence 10, Appl	392	6	2.2	224	4	US-09-393-634-9	Sequence 9, Appl

393	6	2.2	225	4	US-09-540-236-3659	Sequence 3659, Ap	466	6	2.2	310	2	US-08-337-358-42	Sequence 42, Appl
394	6	2.2	226	4	US-09-489-039A-7662	Sequence 7662, Ap	467	6	2.2	310	5	PCT-US95-07577A-42	Sequence 42, Appl
395	6	2.2	226	4	US-09-540-236-3787	Sequence 3787, Ap	468	6	2.2	310	5	PCT-US95-07577A-42	Sequence 42, Appl
396	6	2.2	228	3	US-09-248-335-70	Sequence 70, Appl	469	6	2.2	311	4	US-09-252-991A-1735	Sequence 1735, A
397	6	2.2	228	4	US-09-328-352-7445	Sequence 7445, Ap	470	6	2.2	311	4	US-09-107-532A-7665	Sequence 7665, Ap
398	6	2.2	229	4	US-09-543-681A-6872	Sequence 6872, Ap	471	6	2.2	311	4	US-09-489-039A-9441	Sequence 9441, Ap
399	6	2.2	232	1	US-08-638-911A-33	Sequence 33, Appl	472	6	2.2	311	4	US-09-489-039A-13109	Sequence 13109, A
400	6	2.2	236	1	US-08-442-063A-42	Sequence 42, Appl	473	6	2.2	311	4	US-09-489-039A-13269	Sequence 13269, A
401	6	2.2	237	4	US-09-489-039A-12899	Sequence 12899, A	474	6	2.2	313	4	US-09-489-039A-6623	Sequence 6623, Ap
402	6	2.2	239	4	US-09-325-932A-65	Sequence 65, Appl	475	6	2.2	316	4	US-09-252-991A-29147	Sequence 29147, A
403	6	2.2	244	4	US-09-107-532A-5393	Sequence 5393, Ap	476	6	2.2	316	4	US-09-328-352-5588	Sequence 5588, Ap
404	6	2.2	249	2	US-08-797-689-18	Sequence 18, Appl	477	6	2.2	316	4	US-09-134-000C-5598	Sequence 5598, Ap
405	6	2.2	249	4	US-09-984-186-18	Sequence 18, Appl	478	6	2.2	322	4	US-09-252-991A-22816	Sequence 22816, A
406	6	2.2	251	4	US-09-253-991A-24493	Sequence 24493, A	479	6	2.2	324	4	US-09-489-039A-7962	Sequence 7962, Ap
407	6	2.2	252	4	US-09-253-991A-29152	Sequence 29152, A	480	6	2.2	329	4	US-09-489-039A-13210	Sequence 13210, A
408	6	2.2	253	4	US-09-489-039A-10070	Sequence 10070, A	481	6	2.2	330	4	US-09-252-991A-27332	Sequence 27332, A
409	6	2.2	255	4	US-09-489-039A-11759	Sequence 11759, A	482	6	2.2	330	4	US-09-489-039A-13373	Sequence 13373, Ap
410	6	2.2	256	4	US-09-328-352-7574	Sequence 7574, Ap	483	6	2.2	331	2	US-09-543-681A-5560	Sequence 5560, Ap
411	6	2.2	257	4	US-09-071-035-32	Sequence 32, Appl	484	6	2.2	331	2	US-08-997-060-182	Sequence 182, Ap
412	6	2.2	258	4	US-09-489-039A-12038	Sequence 12038, A	485	6	2.2	331	2	US-08-997-362-182	Sequence 182, Ap
413	6	2.2	260	3	US-08-961-083-32	Sequence 32, Appl	486	6	2.2	331	4	US-09-095-855-182	Sequence 182, Ap
414	6	2.2	260	4	US-09-536-784-32	Sequence 32, Appl	487	6	2.2	331	4	US-09-324-542-182	Sequence 182, Ap
415	6	2.2	261	4	US-09-634-238-420	Sequence 420, Ap	488	6	2.2	332	4	US-09-205-426-182	Sequence 182, Ap
416	6	2.2	263	4	US-09-489-039A-8239	Sequence 8239, Ap	489	6	2.2	333	1	US-09-543-681A-5560	Sequence 5560, Ap
417	6	2.2	264	4	US-09-328-352-6699	Sequence 6699, Ap	490	6	2.2	333	1	US-09-328-352-7378	Sequence 7378, Ap
418	6	2.2	268	4	US-09-543-681A-4676	Sequence 4676, Ap	491	6	2.2	334	3	US-08-975-762-28	Sequence 28, Appl
419	6	2.2	270	4	US-09-252-991A-24772	Sequence 24772, A	492	6	2.2	334	3	US-08-821-324-28	Sequence 28, Appl
420	6	2.2	271	4	US-09-252-991A-17679	Sequence 17679, A	493	6	2.2	334	3	US-09-295-028-28	Sequence 28, Appl
421	6	2.2	271	4	US-09-252-991A-20200	Sequence 20200, A	494	6	2.2	334	3	US-09-106-582-28	Sequence 28, Appl
422	6	2.2	272	4	US-09-489-039A-10410	Sequence 10410, A	495	6	2.2	334	4	US-09-159-469-28	Sequence 28, Appl
423	6	2.2	275	4	US-09-647-324A-26	Sequence 26, Appl	496	6	2.2	334	4	US-09-693-562-28	Sequence 28, Appl
424	6	2.2	275	4	US-09-134-000C-6416	Sequence 6416, Ap	497	6	2.2	338	3	US-08-922-957-1	Sequence 1, Appl
425	6	2.2	277	4	US-09-071-035-30	Sequence 30, Appl	498	6	2.2	338	3	US-08-922-957-3	Sequence 3, Appl
426	6	2.2	278	4	US-09-134-000C-4965	Sequence 4965, Ap	499	6	2.2	339	3	US-09-134-001C-3913	Sequence 3913, Ap
427	6	2.2	279	4	US-09-134-001C-4878	Sequence 4878, Ap	500	6	2.2	340	1	US-07-828-700-8	Sequence 8, Appl
428	6	2.2	279	4	US-09-489-039A-13210	Sequence 13210, A	501	6	2.2	340	1	US-09-489-039A-10379	Sequence 10379, A
429	6	2.2	280	4	US-08-956-171E-5195	Sequence 5195, Ap	502	6	2.2	341	4	US-09-543-681A-4755	Sequence 4755, Ap
430	6	2.2	281	4	US-09-543-681A-4336	Sequence 4336, Ap	503	6	2.2	342	1	US-08-272-919-2	Sequence 2, Appl
431	6	2.2	282	1	US-08-442-063A-45	Sequence 45, Appl	504	6	2.2	342	1	US-08-619-916-2	Sequence 2, Appl
432	6	2.2	282	4	US-09-252-991A-2548	Sequence 2548, A	505	6	2.2	342	4	US-09-134-001C-4190	Sequence 4190, Ap
433	6	2.2	283	4	US-09-107-532A-6360	Sequence 6360, Ap	506	6	2.2	342	4	US-09-252-991A-25420	Sequence 25420, A
434	6	2.2	285	4	US-09-489-039A-12529	Sequence 12529, A	507	6	2.2	342	5	PCT-US95-08542-2	Sequence 2, Appl
435	6	2.2	286	4	US-09-252-991A-20535	Sequence 20535, A	508	6	2.2	346	4	US-09-328-352-6625	Sequence 6625, Ap
436	6	2.2	286	4	US-09-543-681A-7127	Sequence 7127, Ap	509	6	2.2	347	4	US-09-252-991A-27730	Sequence 27730, A
437	6	2.2	288	4	US-09-134-001C-4005	Sequence 4005, Ap	510	6	2.2	347	4	US-09-328-352-7050	Sequence 7050, Ap
438	6	2.2	289	4	US-09-134-000C-4722	Sequence 4722, Ap	511	6	2.2	347	4	US-09-215-418-5	Sequence 5, Appl
439	6	2.2	289	4	US-09-540-236-2015	Sequence 2015, Ap	512	6	2.2	348	4	US-09-489-039A-8632	Sequence 8632, Ap
440	6	2.2	291	4	US-09-252-991A-28572	Sequence 28572, A	513	6	2.2	350	1	US-07-759-568-3	Sequence 3, Appl
441	6	2.2	292	4	US-09-134-001C-3479	Sequence 3479, Ap	514	6	2.2	350	3	US-08-430-286A-8	Sequence 8, Appl
442	6	2.2	292	4	US-09-134-000C-3587	Sequence 3587, Ap	515	6	2.2	350	3	US-09-252-991A-30112	Sequence 30112, A
443	6	2.2	293	4	US-09-252-991A-23350	Sequence 23350, A	516	6	2.2	351	4	US-09-489-039A-7599	Sequence 7599, Ap
444	6	2.2	294	4	US-09-134-000C-4724	Sequence 4724, Ap	517	6	2.2	352	4	US-09-489-039A-8236	Sequence 8236, Ap
445	6	2.2	294	4	US-09-134-000C-4825	Sequence 4825, Ap	518	6	2.2	353	6	5340934-4	Patent No. 5340934
446	6	2.2	296	4	US-09-543-681A-4665	Sequence 4665, Ap	519	6	2.2	356	4	US-09-252-991A-29888	Sequence 29888, A
447	6	2.2	297	4	US-09-252-991A-28075	Sequence 28075, A	520	6	2.2	357	3	US-08-809-286B-3	Sequence 3, Appl
448	6	2.2	298	3	US-08-922-957-4	Sequence 4, Appl	521	6	2.2	357	4	US-09-543-681A-4670	Sequence 4670, Ap
449	6	2.2	298	4	US-09-107-532A-6664	Sequence 6664, Ap	522	6	2.2	358	4	US-09-107-532A-4143	Sequence 4143, Ap
450	6	2.2	300	4	US-09-328-352-6778	Sequence 6778, Ap	523	6	2.2	359	1	US-08-303-338-4	Sequence 4, Appl
451	6	2.2	302	4	US-09-252-991A-24353	Sequence 24353, A	524	6	2.2	359	3	US-08-458-834-4	Sequence 4, Appl
452	6	2.2	303	3	US-08-158-735A-14	Sequence 14, Appl	525	6	2.2	364	4	US-09-134-000C-3748	Sequence 3748, Ap
453	6	2.2	303	3	US-08-158-735A-15	Sequence 15, Appl	526	6	2.2	365	4	US-09-252-991A-20904	Sequence 20904, A
454	6	2.2	303	4	US-09-205-458-818	Sequence 818, Appl	527	6	2.2	366	4	US-08-545-573A-1	Sequence 1, Appl
455	6	2.2	304	4	US-09-328-352-7028	Sequence 7028, Ap	528	6	2.2	370	3	US-08-467-023-97	Sequence 97, Appl
456	6	2.2	305	4	US-09-252-991A-22096	Sequence 22096, A	529	6	2.2	370	4	US-09-252-991A-23112	Sequence 23112, A
457	6	2.2	306	4	US-09-160-975A-2	Sequence 2, Appl	530	6	2.2	370	4	US-09-134-000C-4746	Sequence 4746, Ap
458	6	2.2	306	4	US-09-107-532A-5464	Sequence 5464, Ap	531	6	2.2	372	1	US-08-597-236-8	Sequence 8, Appl
459	6	2.2	307	1	US-08-442-063A-48	Sequence 48, Appl	532	6	2.2	372	1	US-08-746-682A-8	Sequence 8, Appl
460	6	2.2	307	4	US-09-634-238-245	Sequence 245, Appl	533	6	2.2	372	4	US-09-328-352-5436	Sequence 5436, Ap
461	6	2.2	307	4	US-09-107-532A-4626	Sequence 4626, Ap	534	6	2.2	374	4	US-09-252-991A-24475	Sequence 24475, A
462	6	2.2	308	4	US-09-107-532A-5553	Sequence 5553, Ap	535	6	2.2	375	1	US-08-121-714-5	Sequence 5, Appl
463	6	2.2	308	4	US-09-489-039A-7802	Sequence 7802, Ap	536	6	2.2	375	1	US-08-477-108A-5	Sequence 5, Appl
464	6	2.2	309	4	US-09-134-001C-4952	Sequence 4952, Ap	537	6	2.2	375	2	US-08-477-112-5	Sequence 5, Appl
465	6	2.2	310	1	US-08-433-783-42	Sequence 42, Appl	538	6	2.2	375	4	US-09-328-352-6078	Sequence 6078, Ap

539	6	2.2	375	4	US-09-489-039A-11560	Sequence 11560, A	612	6	2.2	437	4	US-09-134-000C-1874	Sequence 3874, Ap
540	6	2.2	375	5	PCT-US93-08322-5	Sequence 5, Appl1	613	6	2.2	438	4	US-09-540-236-2955	Sequence 2555, Ap
541	6	2.2	376	4	US-09-198-452A-995	Sequence 995, A	614	6	2.2	439	1	US-08-333-388-10	Sequence 10, Appl
542	6	2.2	376	4	US-09-252-991A-19776	Sequence 19776, A	615	6	2.2	439	1	US-08-463-694-10	Sequence 10, Appl
543	6	2.2	383	1	US-08-290-448A-78	Sequence 78, Appl	616	6	2.2	439	1	US-08-694-501-10	Sequence 10, Appl
544	6	2.2	383	1	US-08-290-448A-78	Sequence 78, Appl	617	6	2.2	440	4	US-08-484-200-4	Sequence 4, Appl
545	6	2.2	383	1	US-08-175-069A-78	Sequence 78, Appl	618	6	2.2	440	4	US-09-198-452A-44	Sequence 44, Appl
546	6	2.2	383	4	US-08-461-939B-78	Sequence 78, Appl	619	6	2.2	441	4	US-09-107-532A-4625	Sequence 4625, Ap
547	6	2.2	383	4	US-08-461-939B-78	Sequence 78, Appl	620	6	2.2	441	4	US-09-107-532A-4625	Sequence 4625, Ap
548	6	2.2	384	2	US-08-833-226-2	Sequence 2, Appl1	621	6	2.2	445	4	US-09-134-001C-1948	Sequence 6, Appl1
549	6	2.2	386	4	US-08-545-573A-2	Sequence 2, Appl1	622	6	2.2	447	3	US-08-508-761B-6	Sequence 6, Appl1
550	6	2.2	386	4	US-08-545-573A-39	Sequence 39, Appl	623	6	2.2	454	4	US-09-252-991A-30693	Sequence 30693, A
551	6	2.2	387	1	US-08-290-448A-72	Sequence 72, Appl	624	6	2.2	454	1	US-08-621-125-20	Sequence 20, Appl
552	6	2.2	387	1	US-08-290-448A-72	Sequence 72, Appl	625	6	2.2	456	4	US-08-937-155-20	Sequence 20, Appl
553	6	2.2	387	1	US-08-175-069A-72	Sequence 72, Appl	626	6	2.2	456	4	US-08-937-155-20	Sequence 20, Appl
554	6	2.2	387	4	US-08-461-939B-72	Sequence 72, Appl	627	6	2.2	458	4	US-09-252-991A-11884	Sequence 22114, A
555	6	2.2	387	4	US-08-461-939B-72	Sequence 72, Appl	628	6	2.2	463	3	US-08-792-295-1	Sequence 1, Appl1
556	6	2.2	387	1	US-08-290-448A-59	Sequence 59, Appl	629	6	2.2	463	3	US-09-076-432-1	Sequence 1, Appl1
557	6	2.2	391	1	US-08-290-448A-59	Sequence 59, Appl	630	6	2.2	463	4	US-09-198-452A-115	Sequence 115, App
558	6	2.2	391	1	US-08-175-069A-59	Sequence 59, Appl	631	6	2.2	467	4	US-09-134-001C-3579	Sequence 3579, Ap
559	6	2.2	391	4	US-08-461-939B-59	Sequence 59, Appl	632	6	2.2	469	4	US-09-489-039A-13565	Sequence 13565, A
560	6	2.2	391	4	US-08-461-939B-59	Sequence 59, Appl	633	6	2.2	471	3	US-09-160-494-6	Sequence 6, Appl1
561	6	2.2	393	4	US-09-107-532A-4627	Sequence 4627, Ap	634	6	2.2	475	4	US-09-328-352-4942	Sequence 4942, Ap
562	6	2.2	396	4	US-09-540-236-3802	Sequence 3802, Ap	635	6	2.2	479	4	US-09-252-991A-23608	Sequence 23608, A
563	6	2.2	397	1	US-08-290-448A-76	Sequence 76, Appl	636	6	2.2	482	4	US-09-252-991A-23559	Sequence 23559, A
564	6	2.2	397	1	US-08-290-448A-76	Sequence 76, Appl	637	6	2.2	482	4	US-09-252-991A-28368	Sequence 28368, A
565	6	2.2	397	1	US-08-175-069A-76	Sequence 76, Appl	638	6	2.2	483	4	US-09-336-115C-20	Sequence 20, Appl
566	6	2.2	397	4	US-08-461-939B-76	Sequence 76, Appl	639	6	2.2	484	1	US-08-597-236-2	Sequence 2, Appl1
567	6	2.2	397	4	US-08-461-939B-76	Sequence 76, Appl	640	6	2.2	484	1	US-08-746-682A-2	Sequence 2, Appl1
568	6	2.2	398	1	US-08-290-448A-74	Sequence 74, Appl	641	6	2.2	485	4	US-09-252-991A-21841	Sequence 21841, A
569	6	2.2	398	1	US-08-290-448A-74	Sequence 74, Appl	642	6	2.2	486	4	US-09-489-039A-8181	Sequence 8181, Ap
570	6	2.2	398	1	US-08-175-069A-74	Sequence 74, Appl	643	6	2.2	487	4	US-09-252-991A-13102	Sequence 31102, A
571	6	2.2	398	4	US-08-461-939B-74	Sequence 74, Appl	644	6	2.2	489	4	US-09-107-532A-4711	Sequence 4711, Ap
572	6	2.2	398	4	US-08-461-939B-74	Sequence 74, Appl	645	6	2.2	494	3	US-08-984-618-3	Sequence 3, Appl1
573	6	2.2	399	4	US-09-252-991A-31028	Sequence 31028 A	646	6	2.2	498	4	US-09-325-998B-57	Sequence 28182, A
574	6	2.2	400	4	US-09-252-991A-822	Sequence 822, App	647	6	2.2	500	4	US-09-499-302A-5	Sequence 57, Appl
575	6	2.2	401	1	US-08-358-803-7	Sequence 7, Appl1	648	6	2.2	500	4	US-09-499-302A-5	Sequence 5, Appl
576	6	2.2	402	4	US-09-489-039A-9618	Sequence 9618, Ap	649	6	2.2	500	4	US-09-323-998B-58	Sequence 58, Appl
577	6	2.2	404	4	US-09-198-452A-718	Sequence 718, App	650	6	2.2	500	4	US-09-323-998B-59	Sequence 59, Appl
578	6	2.2	404	4	US-09-489-039A-10542	Sequence 10542, A	651	6	2.2	502	3	US-09-382-256-12	Sequence 12, Appl
579	6	2.2	405	2	US-08-222-719-2	Sequence 2, Appl1	652	6	2.2	502	3	US-09-395-115-12	Sequence 12, Appl
580	6	2.2	405	2	US-08-470-925-2	Sequence 2, Appl1	653	6	2.2	502	4	US-08-436-265-12	Sequence 12, Appl
581	6	2.2	405	2	US-08-471-613-2	Sequence 2, Appl1	654	6	2.2	502	4	US-09-679-187-12	Sequence 12, Appl
582	6	2.2	405	4	US-09-107-532A-6414	Sequence 6414, Ap	655	6	2.2	502	4	US-09-499-302A-2	Sequence 2, Appl1
583	6	2.2	405	5	PCT-US93-10443-2	Sequence 2, Appl1	656	6	2.2	502	4	US-09-134-000C-6324	Sequence 4, Appl1
584	6	2.2	407	4	US-09-252-991A-27349	Sequence 27349, A	657	6	2.2	502	4	US-09-323-998B-56	Sequence 56, Appl
585	6	2.2	407	4	US-09-252-991A-29531	Sequence 29531, A	658	6	2.2	503	2	US-08-481-337A-2	Sequence 2, Appl1
586	6	2.2	409	4	US-09-613-303-55	Sequence 55, Appl	659	6	2.2	503	2	US-08-696-268B-2	Sequence 2, Appl1
587	6	2.2	409	4	US-10-267-311-55	Sequence 55, Appl	660	6	2.2	503	2	US-09-382-256-2	Sequence 2, Appl1
588	6	2.2	410	4	US-09-252-991A-30590	Sequence 30590, A	661	6	2.2	503	3	US-09-382-256-2	Sequence 2, Appl1
589	6	2.2	410	4	US-09-198-452A-50	Sequence 50, Appl	662	6	2.2	503	3	US-09-382-256-2	Sequence 2, Appl1
590	6	2.2	412	4	US-09-252-991A-17613	Sequence 17613, A	663	6	2.2	503	4	US-09-436-265-2	Sequence 2, Appl1
591	6	2.2	412	4	US-09-252-991A-25085	Sequence 25085, A	664	6	2.2	503	4	US-09-679-187-2	Sequence 2, Appl1
592	6	2.2	412	4	US-09-134-000C-6666	Sequence 6666, Ap	665	6	2.2	503	5	US-08-448-371A-2	Sequence 2, Appl1
593	6	2.2	413	4	US-09-489-039A-13838	Sequence 13838, A	666	6	2.2	503	5	PCT-US94-11328A-4	Sequence 4, Appl1
594	6	2.2	414	4	US-09-489-039A-12570	Sequence 12570, A	667	6	2.2	503	5	PCT-US95-05467-2	Sequence 2, Appl1
595	6	2.2	417	4	US-09-489-039A-12011	Sequence 12011, A	668	6	2.2	505	1	US-08-146-105-16	Sequence 16, Appl
596	6	2.2	422	2	US-08-663-566A-5	Sequence 5, Appl1	669	6	2.2	505	1	US-08-317-847-14	Sequence 14, Appl
597	6	2.2	422	2	US-08-023-610-5	Sequence 5, Appl1	670	6	2.2	505	4	US-09-543-681A-7151	Sequence 7151, Ap
598	6	2.2	422	2	US-08-288-065A-5	Sequence 5, Appl1	671	6	2.2	505	4	US-09-489-039A-8637	Sequence 8637, Ap
599	6	2.2	422	2	US-08-362-240A-5	Sequence 5, Appl1	672	6	2.2	507	4	US-09-091-097-10	Sequence 10, Appl
600	6	2.2	422	3	US-08-604-372A-3	Sequence 3, Appl1	673	6	2.2	507	4	US-09-091-097-14	Sequence 14, Appl
601	6	2.2	423	5	PCT-US95-10245-5	Sequence 5, Appl1	674	6	2.2	507	4	US-09-252-991A-16830	Sequence 16830, A
602	6	2.2	423	4	US-09-134-001C-4717	Sequence 4717, Ap	675	6	2.2	509	1	US-08-146-105-14	Sequence 14, Appl
603	6	2.2	423	4	US-09-252-991A-21128	Sequence 21128, A	676	6	2.2	509	1	US-08-317-847-14	Sequence 14, Appl
604	6	2.2	424	4	US-09-252-991A-18970	Sequence 18970, A	677	6	2.2	509	2	US-08-481-337A-4	Sequence 4, Appl1
605	6	2.2	425	4	US-09-252-991A-30735	Sequence 30735, A	678	6	2.2	509	3	US-09-382-256-4	Sequence 4, Appl1
606	6	2.2	426	4	US-09-489-039A-9881	Sequence 9881, Ap	679	6	2.2	509	3	US-09-395-115-4	Sequence 4, Appl1
607	6	2.2	429	1	US-08-745-977-4	Sequence 4, Appl1	680	6	2.2	509	4	US-08-123-934A-6	Sequence 6, Appl1
608	6	2.2	429	3	US-09-040-699A-4	Sequence 4, Appl1	681	6	2.2	509	4	US-08-436-265-4	Sequence 4, Appl1
609	6	2.2	431	4	US-09-543-681A-7623	Sequence 7623, Ap	682	6	2.2	509	4	US-09-679-187-4	Sequence 4, Appl1
610	6	2.2	433	4	US-09-252-991A-28924	Sequence 28924, Ap	683	6	2.2	509	4	US-09-874-628-6	Sequence 6, Appl1
611	6	2.2	435	4	US-09-252-991A-28507	Sequence 28507, A	684	6	2.2	509	4	US-08-448-371A-4	Sequence 4, Appl1

685	6	2.2	509	5	PCT-US94-10080-6	Sequence 6, Appl1	758	6	2.2	594	4	US-09-570-383-2	Sequence 2, Appl1
686	6	2.2	509	5	PCT-US95-05467-4	Sequence 4, Appl1	759	6	2.2	594	4	US-09-395-336-8	Sequence 2, Appl1
687	6	2.2	510	4	US-09-252-991A-19584	Sequence 19584, A	760	6	2.2	602	4	US-09-540-236-263	Sequence 3,633, Ap
688	6	2.2	511	4	US-09-134-000C-5362	Sequence 5362, Ap	761	6	2.2	603	4	US-09-543-681A-5587	Sequence 5587, Ap
689	6	2.2	513	4	US-09-489-039A-14224	Sequence 14224, A	762	6	2.2	604	4	US-09-328-352-4437	Sequence 6437, Ap
690	6	2.2	519	4	US-09-252-991A-30815	Sequence 30815, A	763	6	2.2	608	4	US-09-134-000C-4096	Sequence 4096, Ap
691	6	2.2	522	4	US-09-543-681A-7135	Sequence 7135, Ap	764	6	2.2	609	4	US-09-252-991A-32547	Sequence 32547, A
692	6	2.2	525	4	US-09-676-166A-3	Sequence 3, Appl1	765	6	2.2	610	4	US-09-252-991A-27740	Sequence 27740, A
693	6	2.2	525	4	US-09-252-991A-20491	Sequence 20491, A	766	6	2.2	613	4	US-09-252-991A-21411	Sequence 21411, A
694	6	2.2	526	4	US-09-548-938A-10	Sequence 10, Appl1	767	6	2.2	614	3	US-09-017-706-9	Sequence 9, Appl1
695	6	2.2	526	4	US-09-489-039A-8223	Sequence 8223, Ap	768	6	2.2	614	3	US-09-017-706-10	Sequence 11, Appl1
696	6	2.2	527	4	US-08-659-939-4	Sequence 4, Appl1	769	6	2.2	614	3	US-09-017-706-11	Sequence 11, Appl1
697	6	2.2	527	4	US-08-850-041-4	Sequence 4, Appl1	770	6	2.2	614	3	US-09-017-706-12	Sequence 13, Appl1
698	6	2.2	527	4	US-08-896-500-4	Sequence 4, Appl1	771	6	2.2	614	3	US-09-017-706-13	Sequence 14, Appl1
699	6	2.2	527	3	US-08-476-397-4	Sequence 4, Appl1	772	6	2.2	614	3	US-09-567-003C-22	Sequence 22, Appl1
700	6	2.2	527	3	US-08-973-720-4	Sequence 4, Appl1	773	6	2.2	622	1	US-08-547-197-1	Sequence 1, Appl1
701	6	2.2	527	3	US-08-923-558-2	Sequence 4, Appl1	774	6	2.2	622	1	US-08-957-940-1	Sequence 1, Appl1
702	6	2.2	527	3	US-09-262-927-4	Sequence 4, Appl1	775	6	2.2	631	4	US-09-252-991A-18000	Sequence 1000, A
703	6	2.2	527	4	US-09-486-553-2	Sequence 2, Appl1	776	6	2.2	631	4	US-09-640-958-10	Sequence 10, Appl1
704	6	2.2	527	4	US-09-570-383-4	Sequence 4, Appl1	777	6	2.2	634	4	US-09-252-991A-33100	Sequence 33100, A
705	6	2.2	527	4	US-09-395-336-4	Sequence 4, Appl1	778	6	2.2	635	4	US-09-199-452A-128	Sequence 128, Ap
706	6	2.2	529	4	US-09-149-776-732	Sequence 732, App	779	6	2.2	636	4	US-09-134-000C-4594	Sequence 4594, Ap
707	6	2.2	529	4	US-09-328-352-6406	Sequence 6406, Ap	780	6	2.2	637	4	US-09-336-115C-14	Sequence 14, Appl1
708	6	2.2	530	4	US-08-975-762-73	Sequence 73, Appl1	781	6	2.2	640	4	US-09-252-991A-24554	Sequence 24554, A
709	6	2.2	530	3	US-09-295-028-73	Sequence 73, Appl1	782	6	2.2	652	4	US-09-252-991A-17333	Sequence 17333, A
710	6	2.2	530	4	US-09-106-582-73	Sequence 73, Appl1	783	6	2.2	659	4	US-09-124-001C-4303	Sequence 4303, Ap
711	6	2.2	530	4	US-09-159-469-73	Sequence 73, Appl1	784	6	2.2	661	4	US-09-556-916-26	Sequence 26, Appl1
712	6	2.2	530	4	US-09-693-542-73	Sequence 73, Appl1	785	6	2.2	672	4	US-09-556-916-26	Sequence 28, Appl1
713	6	2.2	541	4	US-09-252-991A-20524	Sequence 20524, A	786	6	2.2	672	4	US-09-336-115C-10	Sequence 10, Appl1
714	6	2.2	541	4	US-09-252-991A-32947	Sequence 32947, A	787	6	2.2	672	4	US-09-336-115C-10	Sequence 10, Appl1
715	6	2.2	542	4	US-09-215-694-9	Sequence 9, Appl1	788	6	2.2	674	4	US-09-134-000C-4004	Sequence 4004, Ap
716	6	2.2	544	4	US-09-252-991A-19874	Sequence 19874, A	789	6	2.2	674	4	US-09-252-991A-29071	Sequence 29071, A
717	6	2.2	546	4	US-09-066-047-2	Sequence 2, Appl1	790	6	2.2	681	6	5194595-19	Patent No. 5194595
718	6	2.2	546	4	US-09-489-039A-9627	Sequence 9627, Ap	791	6	2.2	684	4	US-09-252-991A-24802	Sequence 24802, A
719	6	2.2	548	2	US-08-457-822-31	Sequence 31, Appl1	792	6	2.2	684	4	US-09-328-352-5781	Sequence 5781, Ap
720	6	2.2	548	2	US-08-432-697-31	Sequence 31, Appl1	793	6	2.2	684	4	US-09-328-352-5781	Sequence 5781, Ap
721	6	2.2	548	3	US-08-466-248-31	Sequence 31, Appl1	794	6	2.2	700	4	US-09-252-991A-29072	Sequence 29072, A
722	6	2.2	548	4	US-09-149-476-469	Sequence 469, App	795	6	2.2	702	4	US-09-328-352-5781	Sequence 3591, Ap
723	6	2.2	554	4	US-09-564-805-236	Sequence 236, App	796	6	2.2	703	4	US-09-540-236-3691	Sequence 4, Appl1
724	6	2.2	560	1	US-08-052-404-23	Sequence 23, Appl1	797	6	2.2	707	4	US-09-021-560-4	Sequence 4, Appl1
725	6	2.2	560	1	US-08-479-156-23	Sequence 23, Appl1	798	6	2.2	708	4	US-09-336-115C-2	Sequence 2, Appl1
726	6	2.2	561	1	US-08-052-404-24	Sequence 24, Appl1	799	6	2.2	715	4	US-09-134-000C-5094	Sequence 5094, Ap
727	6	2.2	561	1	US-08-479-156-24	Sequence 24, Appl1	800	6	2.2	716	4	US-09-883-533A-2	Sequence 2, Appl1
728	6	2.2	561	3	US-09-227-717-2	Sequence 2, Appl1	801	6	2.2	721	4	US-09-184-748-2	Sequence 2, Appl1
729	6	2.2	562	4	US-09-173-151A-22	Sequence 22, Appl1	802	6	2.2	726	4	US-09-252-991A-26411	Sequence 26411, A
730	6	2.2	567	4	US-09-198-452A-994	Sequence 994, Appl1	803	6	2.2	733	4	US-09-336-115C-4	Sequence 4, Appl1
731	6	2.2	573	4	US-09-252-991A-24488	Sequence 24488, A	804	6	2.2	741	4	US-09-252-991A-20098	Sequence 20098, A
732	6	2.2	573	4	US-09-134-000C-5258	Sequence 5258, Ap	805	6	2.2	744	4	US-09-021-560-2	Sequence 6, Appl1
733	6	2.2	574	4	US-09-252-991A-18912	Sequence 18912, A	806	6	2.2	745	4	US-09-336-115C-6	Sequence 6, Appl1
734	6	2.2	574	4	US-09-252-991A-26497	Sequence 26497, A	807	6	2.2	748	3	US-09-061-764A-19	Sequence 19, Appl1
735	6	2.2	574	6	5223254-2	Patent No. 5223254	808	6	2.2	750	4	US-09-252-991A-26975	Sequence 26975, A
736	6	2.2	577	4	US-09-489-039A-9575	Sequence 9575, A	809	6	2.2	751	4	US-09-966-415-2	Sequence 2, Appl1
737	6	2.2	580	4	US-09-252-991A-22943	Sequence 22943, A	810	6	2.2	754	4	US-08-547-197-2	Sequence 10222, A
738	6	2.2	585	4	US-09-134-001C-4914	Sequence 4914, Ap	811	6	2.2	760	1	US-08-547-197-2	Sequence 2, Appl1
739	6	2.2	590	3	US-08-975-762-64	Sequence 64, Appl1	812	6	2.2	760	1	US-08-957-940-2	Sequence 1816, A
740	6	2.2	590	3	US-09-295-028-64	Sequence 64, Appl1	813	6	2.2	764	4	US-09-252-991A-31816	Sequence 31816, A
741	6	2.2	590	4	US-09-106-582-64	Sequence 64, Appl1	814	6	2.2	769	4	US-09-788-657-15	Sequence 15, Appl1
742	6	2.2	590	4	US-09-159-469-64	Sequence 64, Appl1	815	6	2.2	770	2	US-08-209-521-13	Sequence 13, Appl1
743	6	2.2	590	4	US-09-693-542-64	Sequence 64, Appl1	816	6	2.2	770	3	US-08-961-810-123	Sequence 123, App
744	6	2.2	593	4	US-09-328-352-4866	Sequence 4866, Ap	817	6	2.2	770	3	US-08-352-902D-123	Sequence 123, App
745	6	2.2	594	2	US-08-659-939-2	Sequence 2, Appl1	818	6	2.2	771	4	US-09-265-503B-123	Sequence 123, App
746	6	2.2	594	2	US-08-850-041-2	Sequence 2, Appl1	819	6	2.2	771	4	US-09-252-991A-20455	Sequence 20455, A
747	6	2.2	594	2	US-08-467-563C-6	Sequence 6, Appl1	820	6	2.2	773	4	US-09-252-991A-31403	Sequence 31403, A
748	6	2.2	594	2	US-08-838-189D-6	Sequence 6, Appl1	821	6	2.2	774	4	US-09-252-991A-29487	Sequence 29487, A
749	6	2.2	594	3	US-08-852-344D-6	Sequence 6, Appl1	822	6	2.2	780	1	US-08-186-228-50	Sequence 50, Appl1
750	6	2.2	594	3	US-08-896-500-2	Sequence 2, Appl1	823	6	2.2	780	1	US-08-333-643-14	Sequence 40, Appl1
751	6	2.2	594	3	US-08-476-397-2	Sequence 2, Appl1	824	6	2.2	781	2	US-08-333-643-14	Sequence 50, Appl1
752	6	2.2	594	3	US-08-973-720-2	Sequence 2, Appl1	825	6	2.2	781	2	US-08-675-631-3	Sequence 3, Appl1
753	6	2.2	594	3	US-08-344-639E-6	Sequence 6, Appl1	826	6	2.2	781	2	US-08-352-916A-3	Sequence 3, Appl1
754	6	2.2	594	3	US-08-262-927-2	Sequence 2, Appl1	827	6	2.2	801	3	US-09-248-776-3	Sequence 2, Appl1
755	6	2.2	594	3	US-08-467-569A-6	Sequence 6, Appl1	828	6	2.2	802	3	US-09-104-070-2	Sequence 4, Appl1
756	6	2.2	594	3	US-08-467-561A-6	Sequence 6, Appl1	829	6	2.2	802	3	US-09-147-226-4	Sequence 4, Appl1
757	6	2.2	594	3	US-08-001-554A-6	Sequence 6, Appl1	830	6	2.2	802	4	US-09-522-474-4	Sequence 4, Appl1

831	6	2.2	804	1	US-08-785-428-2	Sequence 2, Appl1	904	6	2.2	1647	4	US-09-535-008-2	Sequence 2, Appl1
832	6	2.2	804	2	US-08-996-797-2	Sequence 2, Appl1	905	6	2.2	1649	4	US-09-535-008-75	Sequence 75, Appl1
833	6	2.2	813	4	US-09-328-352-7421	Sequence 7421, Ap	906	6	2.2	1650	4	US-09-535-008-11	Sequence 11, Appl1
834	6	2.2	819	4	US-09-328-352-5342	Sequence 5342, Ap	907	6	2.2	1650	4	US-09-535-008-11	Sequence 69, Appl1
835	6	2.2	826	5	US-07-638-431-2	Sequence 2, Appl1	908	6	2.2	1678	4	US-09-535-008-55	Sequence 65, Appl1
836	6	2.2	826	5	PCT-US92-00018-2	Sequence 2, Appl1	909	6	2.2	1681	4	US-09-535-008-77	Sequence 77, Appl1
837	6	2.2	826	5	US-09-491-356C-20	Sequence 19, Appl1	910	6	2.2	1682	4	US-09-535-008-73	Sequence 73, Appl1
838	6	2.2	832	4	US-09-585-858-19	Sequence 3, Appl1	911	6	2.2	1690	4	US-09-535-008-39	Sequence 39, Appl1
839	6	2.2	860	1	US-08-117-362-3	Sequence 3, Appl1	912	6	2.2	1784	3	US-08-652-426A-2	Sequence 2, Appl1
840	6	2.2	860	1	US-08-486-924-3	Sequence 3, Appl1	913	6	2.2	1784	3	US-08-652-426A-2	Sequence 2, Appl1
841	6	2.2	860	1	US-08-486-924-3	Sequence 3, Appl1	914	6	2.2	2101	1	US-08-466-390-4	Sequence 4, Appl1
842	6	2.2	871	4	US-09-252-991A-33057	Sequence 33057, A	915	6	2.2	2101	1	US-08-470-950-4	Sequence 4, Appl1
843	6	2.2	873	4	US-09-252-991A-30504	Sequence 30504, A	916	6	2.2	2101	1	US-08-467-781-4	Sequence 4, Appl1
844	6	2.2	879	1	US-08-306-546C-2	Sequence 2, Appl1	917	6	2.2	2101	1	US-08-195-487-4	Sequence 4, Appl1
845	6	2.2	879	2	US-08-530-524A-2	Sequence 2, Appl1	918	6	2.2	2101	2	US-08-483-924-4	Sequence 4, Appl1
846	6	2.2	880	4	US-09-252-991A-25046	Sequence 25046, A	919	6	2.2	2101	3	US-09-452-284-1	Sequence 4, Appl1
847	6	2.2	883	4	US-09-489-039A-12755	Sequence 12755, A	920	6	2.2	2101	5	PCT-US93-06160-4	Sequence 4, Appl1
848	6	2.2	889	4	US-09-252-991A-30096	Sequence 30096, A	921	6	2.2	2142	4	US-09-540-236-1459	Sequence 3459, Ap
849	6	2.2	892	4	US-09-585-858-4	Sequence 4, Appl1	922	6	2.2	2175	4	US-09-404-650-2	Sequence 2, Appl1
850	6	2.2	892	4	US-09-585-858-2	Sequence 2, Appl1	923	6	2.2	2175	4	US-09-404-650-2	Sequence 2, Appl1
851	6	2.2	892	4	US-09-585-858-3	Sequence 3, Appl1	924	6	2.2	2188	4	US-09-404-650-4	Sequence 4, Appl1
852	6	2.2	907	3	US-08-989-299-12	Sequence 12, Appl1	925	6	2.2	2188	4	US-09-404-650-4	Sequence 4, Appl1
853	6	2.2	907	3	US-09-407-427-12	Sequence 4, Appl1	926	6	2.2	2314	4	US-09-268-347-49	Sequence 4, Appl1
854	6	2.2	911	3	US-09-074-579-4	Sequence 4, Appl1	927	6	2.2	2385	4	US-09-343-681A-6304	Sequence 6304, Ap
855	6	2.2	911	3	US-09-388-774-4	Sequence 4, Appl1	928	6	2.2	2431	1	US-07-920-281C-2	Sequence 2, Appl1
856	6	2.2	917	4	US-09-489-039A-12582	Sequence 12582, A	929	6	2.2	2431	3	US-08-466-277-2	Sequence 2, Appl1
857	6	2.2	939	5	PCT-US94-05905-20	Sequence 20, Appl1	930	6	2.2	2464	4	US-09-376-594-726	Sequence 2, Appl1
858	6	2.2	946	4	US-09-328-352-7973	Sequence 7973, Ap	931	6	2.2	2482	1	US-08-328-254-6	Sequence 6, Appl1
859	6	2.2	948	4	US-09-613-303-21	Sequence 21, Appl1	932	6	2.2	2595	3	US-09-306-887A-2	Sequence 2, Appl1
860	6	2.2	948	4	US-10-267-311-21	Sequence 21, Appl1	933	6	2.2	2595	3	US-09-370-700-2	Sequence 2, Appl1
861	6	2.2	951	4	US-09-924-097A-15	Sequence 15, Appl1	934	6	2.2	2595	4	US-09-603-207-2	Sequence 2, Appl1
862	6	2.2	957	4	US-09-252-991A-32568	Sequence 32568, A	935	6	2.2	2842	1	US-07-741-940-7	Sequence 7, Appl1
863	6	2.2	980	4	US-09-252-991A-30838	Sequence 30838, A	936	6	2.2	2842	1	US-08-289-548A-7	Sequence 7, Appl1
864	6	2.2	993	1	US-08-468-537-2	Sequence 2, Appl1	937	6	2.2	2842	1	US-08-452-654-7	Sequence 7, Appl1
865	6	2.2	1009	4	US-09-693-146-4	Sequence 4, Appl1	938	6	2.2	2842	4	US-08-449-731-7	Sequence 7, Appl1
866	6	2.2	1027	3	US-08-446-137B-2	Sequence 2, Appl1	939	6	2.2	2843	1	US-07-741-940-2	Sequence 2, Appl1
867	6	2.2	1032	4	US-09-976-539A-214	Sequence 214, App	940	6	2.2	2843	1	US-08-289-548A-2	Sequence 2, Appl1
868	6	2.2	1075	4	US-09-198-452A-916	Sequence 916, App	941	6	2.2	2843	1	US-08-452-654-2	Sequence 2, Appl1
869	6	2.2	1089	4	US-09-252-991A-20334	Sequence 20334, A	942	6	2.2	2843	1	US-08-452-655B-2	Sequence 2, Appl1
870	6	2.2	1093	3	US-08-545-869D-55	Sequence 55, Appl1	943	6	2.2	2843	1	US-08-452-655B-2	Sequence 2, Appl1
871	6	2.2	1093	3	PCT-US94-04496-55	Sequence 55, Appl1	944	6	2.2	2843	2	US-08-370-235A-2	Sequence 2, Appl1
872	6	2.2	1102	4	US-09-358-383C-36	Sequence 36, Appl1	945	6	2.2	2843	3	US-08-450-582-2	Sequence 2, Appl1
873	6	2.2	1125	4	US-09-328-352-8233	Sequence 8233, Ap	946	6	2.2	2843	3	US-08-450-582-7	Sequence 7, Appl1
874	6	2.2	1140	3	US-09-220-081-2	Sequence 2, Appl1	947	6	2.2	2843	4	US-08-449-731-2	Sequence 4, Appl1
875	6	2.2	1140	3	US-09-677-575-2	Sequence 2, Appl1	948	6	2.2	2873	1	US-08-466-033-15	Sequence 15, Appl1
876	6	2.2	1156	4	US-09-198-452A-171	Sequence 171, App	949	6	2.2	2873	2	US-08-438-911A-2	Sequence 2, Appl1
877	6	2.2	1166	4	US-09-200-650E-7	Sequence 7, Appl1	950	6	2.2	2873	2	US-08-444-733-15	Sequence 15, Appl1
878	6	2.2	1194	2	US-08-680-328-35	Sequence 35, Appl1	951	6	2.2	2873	2	US-08-464-134-15	Sequence 15, Appl1
879	6	2.2	1228	4	US-09-439-313-537	Sequence 537, App	952	6	2.2	2873	2	US-08-461-361-15	Sequence 15, Appl1
880	6	2.2	1228	4	US-09-636-215-537	Sequence 537, App	953	6	2.2	2873	2	US-08-461-361-15	Sequence 15, Appl1
881	6	2.2	1228	4	US-09-685-166A-537	Sequence 537, App	954	6	2.2	2873	5	PCT-US95-06266-15	Sequence 15, Appl1
882	6	2.2	1245	2	US-08-801-263A-3	Sequence 3, Appl1	955	6	2.2	2905	4	US-08-469-260A-401	Sequence 401, App
883	6	2.2	1245	2	US-08-801-263A-6	Sequence 6, Appl1	956	6	2.2	2905	4	US-08-488-446-401	Sequence 401, App
884	6	2.2	1245	2	US-08-801-263A-10	Sequence 10, Appl1	957	6	2.2	2905	4	US-08-467-344A-401	Sequence 401, App
885	6	2.2	1245	3	US-09-102-248-3	Sequence 3, Appl1	958	6	2.2	2910	1	US-08-466-033-183	Sequence 183, App
886	6	2.2	1245	3	US-09-102-248-6	Sequence 6, Appl1	959	6	2.2	2910	2	US-08-444-733-183	Sequence 183, App
887	6	2.2	1245	3	US-09-102-248-10	Sequence 10, Appl1	960	6	2.2	2910	2	US-08-464-134-183	Sequence 183, App
888	6	2.2	1245	3	US-09-367-764-3	Sequence 3, Appl1	961	6	2.2	2910	2	US-08-461-361-183	Sequence 183, App
889	6	2.2	1245	4	US-09-367-764-6	Sequence 6, Appl1	962	6	2.2	2910	2	US-08-461-361-183	Sequence 183, App
890	6	2.2	1245	4	US-09-367-764-10	Sequence 10, Appl1	963	6	2.2	2910	2	US-08-461-361-183	Sequence 183, App
891	6	2.2	1261	4	US-09-439-313-538	Sequence 538, App	964	6	2.2	2910	2	PCT-US95-06266-157	Sequence 157, App
892	6	2.2	1261	4	US-09-439-313-538	Sequence 538, App	965	6	2.2	2910	2	US-08-821-355A-7	Sequence 7, Appl1
893	6	2.2	1261	4	US-09-685-215-538	Sequence 538, App	966	6	2.2	2910	2	US-09-003-687A-7	Sequence 7, Appl1
894	6	2.2	1285	1	US-07-583-945-2	Sequence 2, Appl1	967	6	2.2	3021	1	US-07-689-008-2	Sequence 2, Appl1
895	6	2.2	1285	1	US-08-453-141-2	Sequence 2, Appl1	968	6	2.2	3021	1	US-09-413-814-93	Sequence 93, Appl1
896	6	2.2	1285	3	US-08-293-314-2	Sequence 3, Appl1	969	6	2.2	3079	3	US-09-413-814-80	Sequence 80, Appl1
897	6	2.2	1297	4	US-09-328-352-6373	Sequence 6373, Ap	970	6	2.2	3168	4	US-09-489-039A-44067	Sequence 14067, A
898	6	2.2	1341	4	US-09-252-991A-26785	Sequence 26785, A	971	6	2.2	3170	2	US-07-642-734C-5	Sequence 5, Appl1
899	6	2.2	1374	4	US-09-252-991A-24636	Sequence 24636, A	972	6	2.2	3201	4	US-08-439-009A-5	Sequence 5, Appl1
900	6	2.2	1427	4	US-09-252-991A-27005	Sequence 27005, A	973	6	2.2	3248	5	US-09-679-279-15	Sequence 15, Appl1
901	6	2.2	1461	4	US-09-976-594-531	Sequence 531, App	974	6	2.2	3348	5	US-08-353-700-1	Sequence 1, Appl1
902	6	2.2	1487	4	US-09-489-039A-12113	Sequence 12113, A	975	6	2.2	3491	2	PCT-US95-16216-1	Sequence 1, Appl1
903	6	2.2	1646	4	US-09-535-008-67	Sequence 67, Appl1	976	6	2.2	3491	2	US-07-642-734C-2	Sequence 2, Appl1

```

977 6 2.2 3491 3 US-08-439-009A-2 Sequence 2, Appl
978 6 2.2 3546 4 US-09-679-279-13 Sequence 13, Appl
979 6 2.2 3898 4 US-08-750-717-2 Sequence 2, Appl
980 5 1.8 5 4 US-08-877-605-330 Sequence 330, Appl
981 5 1.8 6 1 US-08-704-170-81 Sequence 81, Appl
982 5 1.8 6 5 PCT-US94-02631-81 Sequence 81, Appl
983 5 1.8 7 1 US-08-346-333-83 Sequence 83, Appl
984 5 1.8 7 3 US-09-256-754-389 Sequence 389, Appl
985 5 1.8 7 3 US-09-042-107-389 Sequence 389, Appl
986 5 1.8 7 3 US-08-667-530-63 Sequence 63, Appl
987 5 1.8 7 3 US-09-457-0468-40 Sequence 40, Appl
988 5 1.8 7 4 US-09-731-558-19 Sequence 19, Appl
989 5 1.8 7 4 US-09-989-789-159 Sequence 159, Appl
990 5 1.8 7 4 US-09-989-789-275 Sequence 275, Appl
991 5 1.8 7 4 US-09-989-789-378 Sequence 378, Appl
992 5 1.8 7 4 US-09-989-789-383 Sequence 383, Appl
993 5 1.8 7 4 US-09-989-789-437 Sequence 437, Appl
994 5 1.8 7 4 US-09-989-789-633 Sequence 633, Appl
995 5 1.8 7 4 US-09-989-789-664 Sequence 664, Appl
996 5 1.8 7 4 US-09-989-789-760 Sequence 760, Appl
997 5 1.8 7 4 US-09-989-789-1077 Sequence 1077, Appl
998 5 1.8 7 4 US-09-989-789-1078 Sequence 1078, Appl
999 5 1.8 7 4 US-09-989-789-1761 Sequence 1761, Appl
1000 5 1.8 7 4 US-09-989-789-1765 Sequence 1765, Appl

```

ALIGNMENTS

```

RESULT 1
US-09-540-236-2991
; Sequence 2991, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540.236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2991
; LENGTH: 118
; TYPE: PRT
; ORGANISM: M.catarhalis
US-09-540-236-2991

```

```

Query Match 20.7%; Score 57; DB 4; Length 118;
Best Local Similarity 100.0%; Pred. No. 2.2e-47;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 47 PEOAAVEYAGVAKEXNLTVELEFNDYAMPNSAVSKGLDANAMQHRYLEKDSQ 103
DB 57 PEOAAVEYAGVAKEXNLTVELEFNDYAMPNSAVSKGLDANAMQHRYLEKDSQ 113

```

```

RESULT 2
US-09-543-681A-6534
; Sequence 6534, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILI
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543.681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 6534
; LENGTH: 279
; TYPE: PRT

```

```

; ORGANISM: Proteus mirabilis
US-09-543-681A-6534
Query Match 5.8%; Score 16; DB 4; Length 279;
Best Local Similarity 100.0%; Pred. No. 2e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 224 FVEDKDSFYNNITVAR 239
DB 227 FVEDKDSFYNNITVAR 242

RESULT 3
US-09-328-352-6702
; Sequence 6702, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-033A
; CURRENT APPLICATION NUMBER: US/09/328.352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6702
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6702

```

```

Query Match 3.6%; Score 10; DB 4; Length 286;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 195 ARAIDVDVLA 204
DB 200 ARAIDVDVLA 209

```

```

RESULT 4
US-09-489-039A-10641
; Sequence 10641, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489.039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10641
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10641

```

```

Query Match 3.3%; Score 9; DB 4; Length 285;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 133 ELXDGAATIA 141
DB 140 ELXDGAATIA 148

```

```

RESULT 5
US-09-107-532A-7234
; Sequence 7234, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:

```

APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
FILING DATE: 30-Jun-1998
APPLICATION NUMBER: US/09/107,532A
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 7234:
SEQUENCE CHARACTERISTICS:
LENGTH: 291 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc.feature
LOCATION: (B) LOCATION 1...291
SEQUENCE DESCRIPTION: SEQ ID NO: 7234:
US-09-107-532A-7234

Query Match 3.3%; Score 9; DB 4; Length 291;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 PYNNIVAR 239
Db 240 PYNNIVAR 248

RESULT 6
US-09-328-352-6924
Sequence 6924, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: BADVANNIT FOR DIAGNOSTICS AND THERAPEUTICS
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 6924
LENGTH: 286
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-6924

Query Match 3.3%; Score 9; DB 4; Length 296;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 148 NLARLIL 156
Db 164 NLARLIL 172

RESULT 7
US-09-107-532A-6443
Sequence 6443, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 6443:
SEQUENCE CHARACTERISTICS:
LENGTH: 269 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc.feature
LOCATION: (B) LOCATION 1...269
SEQUENCE DESCRIPTION: SEQ ID NO: 6443:
US-09-107-532A-6443

Query Match 2.9%; Score 8; DB 4; Length 269;
Best Local Similarity 100.0%; Pred. No. 1.0;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 ELKDGAT 140
Db 118 ELKDGAT 125

RESULT 8
US-09-252-991A-19629
Sequence 19629, Application US/09252991A
Patent No. 6551795

```

; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19629
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19629

Query Match      2.9%; Score 8; DB 4; Length 273;
Best Local Similarity 100.0%; Pred. No. 10;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      51 VAEVAGGV 58
      |||||
Db      98 VAEVAGGV 105

RESULT 9
US-09-543-681A-4548
; Sequence 4548, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4548
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-4548

Query Match      2.9%; Score 8; DB 4; Length 279;
Best Local Similarity 100.0%; Pred. No. 11;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      122 AGYSTKIK 129
      |||||
Db      123 AGYSTKIK 130

RESULT 10
US-09-543-681A-6914
; Sequence 6914, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 6914
; LENGTH: 279
; TYPE: PRT

; ORGANISM: Proteus mirabilis
US-09-543-681A-6914

Query Match      2.9%; Score 8; DB 4; Length 279;
Best Local Similarity 100.0%; Pred. No. 11;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      159 QGLIKKD 166
      |||||
Db      160 QGLIKKD 167

RESULT 11
US-09-489-039A-12337
; Sequence 12337, Application US/09489039A
; Patent No. 6610636
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12337
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12337

Query Match      2.9%; Score 8; DB 4; Length 289;
Best Local Similarity 100.0%; Pred. No. 11;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      113 VGNTFVVP 120
      |||||
Db      124 VGNTFVVP 131

RESULT 12
US-09-252-991A-17635
; Sequence 17635, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17635
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17635

Query Match      2.9%; Score 8; DB 4; Length 465;
Best Local Similarity 100.0%; Pred. No. 17;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      149 LARALILL 156
      |||||
Db      95 LARALILL 102

RESULT 13
```

```
US-08-858-207A-422
; Sequence 422, Application US/08858207A
; Patent No. 6348328
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Nicholas, Richard
; APPLICANT: Stodola, Robert
; TITLE OF INVENTION: No. 6348328el Compounds
; NUMBER OF SEQUENCES: 552
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/858,207A
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/017670
; FILING DATE: 14-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmil, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50475
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 422:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6348328e
; US-08-858-207A-422

Query Match      2.5%; Score 7; DB 4; Length 78;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      101 DSOEKL 107
DB      34 DSOEKL 40

RESULT 14
US-09-134-001C-3331
; Sequence 3331, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3331
```

```
LENGTH: 116
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3331

Query Match      2.5%; Score 7; DB 4; Length 116;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      65 LTVELVE 71
DB      31 LTVELVE 37

RESULT 15
US-09-963-791-18
; Sequence 18, Application US/09963791
; Patent No. 6649399
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Scoville, John
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6649399el Human Proteases and Polynucleotides Encoding the San
; FILE REFERENCE: LEX-0105-USA
; CURRENT APPLICATION NUMBER: US/09/963,791
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: US 60/169,769
; PRIOR FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-963-791-18

Query Match      2.5%; Score 7; DB 4; Length 159;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      233 VNIIVAR 239
DB      144 VNIIVAR 150

RESULT 16
US-08-572-447C-11
; Sequence 11, Application US/08572447C
; Patent No. 5955090
; GENERAL INFORMATION:
; APPLICANT: Knaap, Bernhard
; APPLICANT: Hungerer, Klaus-Dieter
; APPLICANT: Broxer, Michael
; APPLICANT: Von Specht, Bernd-Ulrich
; APPLICANT: Domdey, Horst
; TITLE OF INVENTION: Immunogenic Hybrid Protein Ogrf-Opti
; TITLE OF INVENTION: Derived From Pseudomonas aeruginosa Membrane Proteins.
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/572,447C
FILING DATE: 14-DEC-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 94120023.0
FILING DATE: 16-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: M. Paul Barker
REGISTRATION NUMBER: 32,013
REFERENCE/DOCKET NUMBER: 05552.1395-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4400
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 161 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-572-447C-11

Query Match 2.5%; Score 7; DB 2; Length 161;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 259 EVEAEAK 265
|||
Db 155 EVEAEAK 161

RESULT 17
US-09-267-747-11
Sequence 11, Application US/09267747
Patent No. 6300102
GENERAL INFORMATION:
APPLICANT: Knapp, Bernhard
APPLICANT: Hungerer, Klaus-Dieter
APPLICANT: Broker, Michael
APPLICANT: Von Specht, Bernd-Ulrich
APPLICANT: Domdey, Horst
TITLE OF INVENTION: Immunogenic Hybrid Protein Opr-Opr
TITLE OF INVENTION: Derived From Pseudomonas aeruginosa Membrane Proteins.
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESS: Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/267,747
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/572,447
FILING DATE: 14-DEC-1995
APPLICATION NUMBER: EP 94120023.0
FILING DATE: 16-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: M. Paul Barker
REGISTRATION NUMBER: 32,013
REFERENCE/DOCKET NUMBER: 05552.1395-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000

TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 161 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-267-747-11

Query Match 2.5%; Score 7; DB 4; Length 161;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 259 EVEAEAK 265
|||
Db 155 EVEAEAK 161

RESULT 18
PCT-US94-02889-2
Sequence 2, Application PC/TUS9402889
GENERAL INFORMATION:
APPLICANT: Lefebvre Dr., Rance B.
APPLICANT: Pering Dr., Guey-Chuen
TITLE OF INVENTION: Methods, Compositions, and
TITLE OF INVENTION: Kits for Diagnosing Lyme
TITLE OF INVENTION: Disease
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: ROBBINS, BERLINER & CARSON
STREET: 201 North Figueroa Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90012
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS 5.0
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/02889
FILING DATE: herewith
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Berliner, Robert
REGISTRATION NUMBER: 20,191
REFERENCE/DOCKET NUMBER: 5555-211-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 977-1001
TELEFAX: (213) 977-1003
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 199 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-02889-2

Query Match 2.5%; Score 7; DB 5; Length 199;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 62 KYNLIVE 68
|||
Db 16 KYNLIVE 22

RESULT 19
US-08-572-447C-15
Sequence 15, Application US/08572447C
Patent No. 555060

GENERAL INFORMATION:
APPLICANT: Knapp, Bernhard
APPLICANT: Hungerer, Klaus-Dieter
APPLICANT: Broker, Michael
APPLICANT: Von Specht, Bernd-Ulrich
APPLICANT: Domdey, Horst
TITLE OF INVENTION: Immunogenic Hybrid Protein Oprf-OprI
TITLE OF INVENTION: Derived From Pseudomonas aeruginosa Membrane Proteins.
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 14-DEC-1995
APPLICATION NUMBER: US/08/572,447C
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 94120023.0
FILING DATE: 16-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: M. Paul Barker
REGISTRATION NUMBER: 32,013
REFERENCE/DOCKET NUMBER: 05552.1395-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4000
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 226 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-572-447C-15

Query Match 2.5%; Score 7; DB 2; Length 226;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 259 EVEAEAK 265
DB 220 EVEAEAK 226

RESULT 20
US-09-267-747-15
Sequence 15, Application US/09267747
Patent No. 6300102
GENERAL INFORMATION:
APPLICANT: Knapp, Bernhard
APPLICANT: Hungerer, Klaus-Dieter
APPLICANT: Broker, Michael
APPLICANT: Von Specht, Bernd-Ulrich
APPLICANT: Domdey, Horst
TITLE OF INVENTION: Immunogenic Hybrid Protein Oprf-OprI
TITLE OF INVENTION: Derived From Pseudomonas aeruginosa Membrane Proteins.
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA

ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 14-DEC-1995
APPLICATION NUMBER: US/09/267,747
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/572,447
FILING DATE: 14-DEC-1995
APPLICATION NUMBER: EP 94120023.0
FILING DATE: 16-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: M. Paul Barker
REGISTRATION NUMBER: 32,013
REFERENCE/DOCKET NUMBER: 05552.1395-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4000
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 226 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-267-747-15

Query Match 2.5%; Score 7; DB 4; Length 226;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 259 EVEAEAK 265
DB 220 EVEAEAK 226

RESULT 21
US-09-543-681A-4708
Sequence 4708, Application US/09543681A
Patent No. 6605709
GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543,681A
PRIOR FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 4708
LENGTH: 274
TYPE: PRT
ORGANISM: Proteus mirabilis
US-09-543-681A-4708

Query Match 2.5%; Score 7; DB 4; Length 274;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 149 IARALIL 155
DB 166 IARALIL 172

RESULT 22
US-09-489-039A-7722
Sequence 7722, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:


```
APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7722
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7722

Query Match      2.5%; Score 7; DB 4; Length 275;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      149 LARATL 155
DB      167 LARATL 173

RESULT 23
US-09-434-354-48
; Sequence 48, Application US/09434354
; Patent No. 6562563
; GENERAL INFORMATION:
; APPLICANT: Murphy, Anne N.
; APPLICANT: Clevenger, William
; APPLICANT: Wiley, Sandra Eileen
; APPLICANT: Andreyev, Alexander Y.
; APPLICANT: Frigeri, Luciano G.
; APPLICANT: Velicelcib, Gonul
; APPLICANT: Davis, Robert E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETERMINING
; TITLE OF INVENTION: INTERACTIONS OF MITOCHONDRIAL COMPONENTS, AND FOR
; TITLE OF INVENTION: IDENTIFYING AGENTS THAT ALTER SUCH INTERACTIONS
; FILE REFERENCE: 660088.433
; CURRENT APPLICATION NUMBER: US/09/434,354
; CURRENT FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 48
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-434-354-48

Query Match      2.5%; Score 7; DB 4; Length 298;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      29 AAISXTA 35
DB      19 AAISXTA 25

RESULT 24
US-09-434-354-49
; Sequence 49, Application US/09434354
; Patent No. 6562563
; GENERAL INFORMATION:
; APPLICANT: Murphy, Anne N.
; APPLICANT: Clevenger, William
; APPLICANT: Wiley, Sandra Eileen
; APPLICANT: Andreyev, Alexander Y.
; APPLICANT: Frigeri, Luciano G.
; APPLICANT: Velicelcib, Gonul
; APPLICANT: Davis, Robert E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETERMINING
; TITLE OF INVENTION: INTERACTIONS OF MITOCHONDRIAL COMPONENTS, AND FOR
; TITLE OF INVENTION: INTERACTIONS OF MITOCHONDRIAL COMPONENTS, AND FOR
```

```
; TITLE OF INVENTION: IDENTIFYING AGENTS THAT ALTER SUCH INTERACTIONS
; FILE REFERENCE: 660088.433
; CURRENT APPLICATION NUMBER: US/09/434,354
; CURRENT FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 49
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-434-354-49

Query Match      2.5%; Score 7; DB 4; Length 298;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      29 AAISXTA 35
DB      19 AAISXTA 25

RESULT 25
US-09-134-001C-3171
; Sequence 3171, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3171
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3171

Query Match      2.5%; Score 7; DB 4; Length 302;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      182 KKVYKE 188
DB      3 KKVYKE 9

RESULT 26
US-09-252-991A-18873
; Sequence 18873, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18873
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18873
```

Query Match 2.5%; Score 7; DB 4; Length 306;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 229 DSPYVNI 235
Db 259 DSPYVNI 265

RESULT 27
US-09-522-714-10
; Sequence 10, Application US/09522714
; Patent No. 6563020
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; APPLICANT: Yalpani, Nasser
; TITLE OF INVENTION: Maize Chitinases and Their Use in
; FILE REFERENCE: 1100
; CURRENT APPLICATION NUMBER: US/09/522,714
; EARLIER FILING DATE: 2000-03-10
; EARLIER FILING DATE: 1999-03-24
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Zea mays
US-09-522-714-10

Query Match 2.5%; Score 7; DB 4; Length 309;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 214 VGLTASE 220
Db 254 VGLTASE 260

RESULT 28
US-09-963-791-8
; Sequence 8, Application US/0963791
; Patent No. 6649399
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Scoville, John
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6649399 Human Proteases and Polynucleotides Encoding the sa
; FILE REFERENCE: LEX-0105-USA
; CURRENT APPLICATION NUMBER: US/09/963,791
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: US 60/169,769
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-963-791-8

Query Match 2.5%; Score 7; DB 4; Length 310;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 233 VNIIVAR 239
Db 295 VNIIVAR 301

RESULT 29
US-09-963-791-16
; Sequence 16, Application US/09963791
; Patent No. 6649399
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Scoville, John
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6649399 Human Proteases and Polynucleotides Encoding the San
; FILE REFERENCE: LEX-0105-USA
; CURRENT APPLICATION NUMBER: US/09/963,791
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: US 60/169,769
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 317
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-963-791-16

Query Match 2.5%; Score 7; DB 4; Length 317;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 233 VNIIVAR 239
Db 144 VNIIVAR 150

RESULT 30
US-09-489-039A-9512
; Sequence 9512, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9512
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9512

Query Match 2.5%; Score 7; DB 4; Length 326;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 DIVENPK 182
Db 222 DIVENPK 228

RESULT 31
US-09-540-236-2111
; Sequence 2111, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARI
; FILE REFERENCE: FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 2709.2005-001
CURRENT APPLICATION NUMBER: US/09/540,236
CURRENT FILING DATE: 2000-04-04
NUMBER OF SEQ ID NOS: 3840
SEQ ID NO 2111
LENGTH: 342
TYPE: PRT
ORGANISM: M.catarhalls
US-09-540-236-2111

Query Match 2.5%; Score 7; DB 4; Length 342;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 149 LARALIL 155
DB 1 LARALIL 7

RESULT 32
US-09-543-681A-5961
Sequence 5961, Application US/09543681A
Patent No. 6605709
GENERAL INFORMATION:
APPLICANT: GARY BRETTON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543,681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 5961
LENGTH: 347
TYPE: PRT
ORGANISM: Proteus mirabilis
US-09-543-681A-5961

Query Match 2.5%; Score 7; DB 4; Length 347;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 149 LARALIL 155
DB 148 LARALIL 154

RESULT 33
US-09-252-991A-30094
Sequence 30094, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 30094
LENGTH: 351
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30094

Query Match 2.5%; Score 7; DB 4; Length 351;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 259 EVEAEAK 265
DB 345 EVEAEAK 351

RESULT 34
US-09-963-791-20
Sequence 20, Application US/09963791
Patent No. 6649399
GENERAL INFORMATION:
APPLICANT: Donoho, Gregory
APPLICANT: Turner, C. Alexander Jr.
APPLICANT: Friedrich, Glenn
APPLICANT: Scoville, John
APPLICANT: Zambrowicz, Brian
APPLICANT: Sands, Arthur T.
TITLE OF INVENTION: No. 6649399el Human Proteases and Polynucleotides Encoding the San
FILE REFERENCE: LEX-0103-USA
CURRENT APPLICATION NUMBER: US/09/963,791
CURRENT FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: US 60/169,769
PRIOR FILING DATE: 1999-12-09
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 20
LENGTH: 356
TYPE: PRT
ORGANISM: Homo sapiens
US-09-963-791-20

Query Match 2.5%; Score 7; DB 4; Length 356;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 233 VNIIVAR 239
DB 144 VNIIVAR 150

RESULT 35
US-09-543-681A-5390
Sequence 5390, Application US/09543681A
Patent No. 6605709
GENERAL INFORMATION:
APPLICANT: GARY BRETTON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543,681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 5390
LENGTH: 361
TYPE: PRT
ORGANISM: Proteus mirabilis
US-09-543-681A-5390

Query Match 2.5%; Score 7; DB 4; Length 361;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 261 EBAEAKQ 267
DB 174 EBAEAKQ 180

RESULT 36
US-08-305-172B-4
Sequence 4, Application US/08305172B
Patent No. 5656470

GENERAL INFORMATION:
APPLICANT: Martinis, Susan A.
APPLICANT: Zhang, Jiansu
APPLICANT: Schimmel, Paul R.
TITLE OF INVENTION: Recombinant Mycobacterial Seryl-tRNA
TITLE OF INVENTION: Synthetase Genes, Tester Strains and Assays
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSES: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Willetta Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: US
ZIP: 02173-4799
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/305,172B
FILING DATE: 13-SEP-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 419 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-305-172B-4

Query Match 2.5%; Score 7; DB 1; Length 419;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 258 DEVEAEA 264
DB 87 DEVEAEA 93

RESULT 37
US-09-963-791-22
Sequence 22, Application US/09963791
Patent No. 6649399
GENERAL INFORMATION:
APPLICANT: Donoho, Gregory
APPLICANT: Turner, C. Alexander Jr.
APPLICANT: Friedrich, Glenn
APPLICANT: Scoville, John
APPLICANT: Zambrowicz, Brian
APPLICANT: Sands, Arthur T.
TITLE OF INVENTION: No. 6649399el Human Proteases and Polynucleotides Encoding the Sa
FILE REFERENCE: LEX-0105-USA
CURRENT APPLICATION NUMBER: US/09/963,791
CURRENT FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: US 60/169,769
PRIOR FILING DATE: 1999-12-09
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 22
LENGTH: 438
TYPE: PRT
ORGANISM: Homo sapiens
US-09-963-791-22

Query Match 2.5%; Score 7; DB 4; Length 438;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 233 VNIIVAR 239
DB 144 VNIIVAR 150

RESULT 38
US-09-252-991A-20116
Sequence 20116, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 20116
TYPE: PRT
LENGTH: 447
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20116

Query Match 2.5%; Score 7; DB 4; Length 447;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 LASGIAL 18
DB 1 LASGIAL 7

RESULT 39
US-09-963-791-6
Sequence 6, Application US/09963791
Patent No. 6649399
GENERAL INFORMATION:
APPLICANT: Donoho, Gregory
APPLICANT: Turner, C. Alexander Jr.
APPLICANT: Friedrich, Glenn
APPLICANT: Scoville, John
APPLICANT: Zambrowicz, Brian
APPLICANT: Sands, Arthur T.
TITLE OF INVENTION: No. 6649399el Human Proteases and Polynucleotides Encoding the San
FILE REFERENCE: LEX-0105-USA
CURRENT APPLICATION NUMBER: US/09/963,791
CURRENT FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: US 60/169,769
PRIOR FILING DATE: 1999-12-09
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 468
TYPE: PRT
ORGANISM: Homo sapiens
US-09-963-791-6

Query Match 2.5%; Score 7; DB 4; Length 468;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 233 VNIIVAR 239
DB 295 VNIIVAR 301

RESULT 40

US-07-686-591-4
; Sequence 4, Application US/07686591
; Patent No. 5215915
; GENERAL INFORMATION:
; APPLICANT: Tiberi, Mario
; APPLICANT: Jarvie, Keith R.
; APPLICANT: Caron, Marc G.
; TITLE OF INVENTION: Cloned Gene Encoding Rat D1B Dopamine Receptor
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenneth D. Sibley; Bell, Seltzer, Park and Gibson
; STREET: Post Office Drawer 34009
; CITY: Charlotte
; STATE: No. 5215915th Carolina
; COUNTRY: U.S.A.
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/686,591
; FILING DATE: 19910406
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sibley, Kenneth D.
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5405.24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; TELEX: 575102
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 475 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-686-591-4

Query Match 2.5%; Score 7; DB 1; Length 475;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 AVAEVAG 56
|||
Db 97 AVAEVAG 103

Search completed: June 16, 2004, 11:18:09
Job time : 31 secs